FEB 10 1921

Vol. VI, No. 5 PSYCHOLOGICAL REVIEW PUBLICATIONS October, 1920

Journal of Experimental Psychology

EDITED BY

JOHN B. WATSON, New York City

HOWARD C. WARREN, Princeton University (*Review*)

JAMES R. ANGELL, New York City (*Monographs*)

SHEPHERD I. FRANZ, Govt. Hosp. for Insane (*Bulletin*) and

MADISON BENTLEY, University of Illinois (*Index*)

ADVISORY EDITORS

R. A. [illegible], Columbia; RAYMOND DODGE, Wesleyan University; KNIGHT DUNLAP, [illegible] University; H. S. LANGFELD, Harvard University; E. L. THORNDIKE, [illegible] University

CONTENTS

The Effect of [illegible] Association, John J. B. Morgan, 319.

Effects of [illegible] Mental and Motor Efficiency: Sven Froeberg, 334.

A Study [illegible] Spatial References to the Lookout and Signal Service, [illegible] G. Rand, and D. Buckley, 347.

The [illegible], C. [illegible], 357.

[illegible] Psychological Laboratory of McLean Hospital: F. L. Wells.

[illegible] Associations: Harold E. Burtt, 390.

[illegible] Tests: Knight Dunlap and Agnes Snyder, 396.

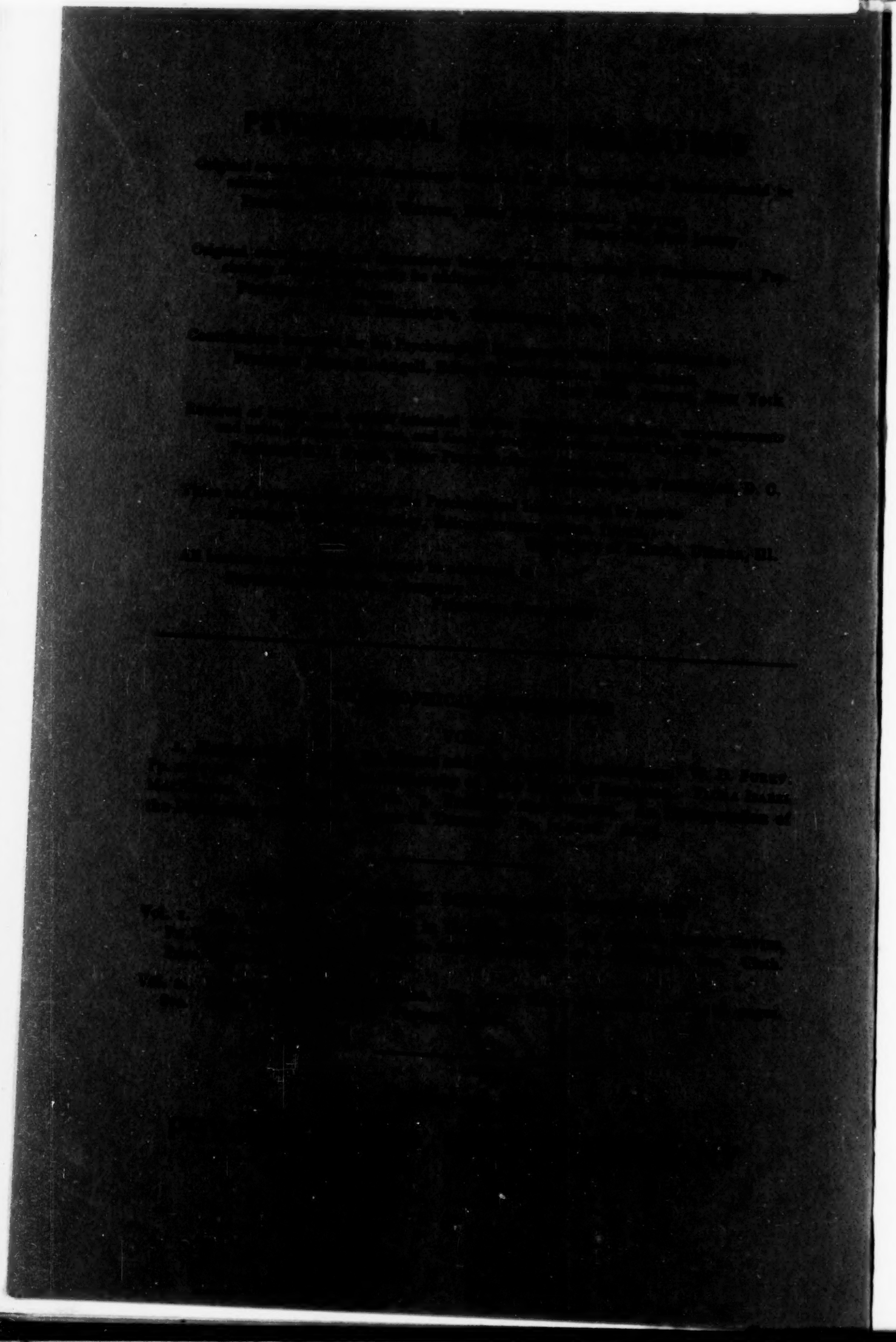
PUBLISHED BY

PSYCHOLOGICAL REVIEW COMPANY

[illegible], LANCASTER, PA., AND

[illegible]

Entered as second-class matter at the post-office at Lancaster, Pa., pending



Journal of
Experimental Psychology

VOL. III, No. 5. OCTOBER, 1920

THE EFFECT OF FATIGUE ON RETENTION

BY JOHN J. B. MORGAN
University of Minnesota

INTRODUCTION

When it is desired to demonstrate muscular fatigue all that one has to do is to isolate a muscle or a group of muscles, attach a load, provide successive stimuli of sufficient intensity and at sufficiently short intervals and ordinary observation will detect progressive weakening in muscular contraction. The fatigue so produced is caused by the catabolism in the muscle and recovery takes place when the fatigue products of this catabolism are removed. The fact that when one does mental work he feels fatigued and desires rest just as one does when he becomes physically fatigued has led scientists to search for experimental evidence of neural fatigue by methods similar to those used to demonstrate muscular fatigue. The failure of such experiments to demonstrate neural fatigue has led to two diverse results. The first was to stimulate investigators to refine their methods and to invent more delicate instruments to detect catabolism in the nerve. The second was to infer that the neural mechanism does not become fatigued, and to assume that the feeling of fatigue after so-called mental work was due either to fatigue of a peripheral sort—that is, to muscles or sense organs—or to a feeling of monotony.

Experiments have shown beyond a doubt that the nerve fibre is much less susceptible to fatigue than the muscle. If the refractory phase is regarded as a fatigue phenomenon, it must be admitted that neural fatigue is relatively slight

and recovery extremely rapid. It is quite possible that this refractory phase forms a resting period and thus gives the nerve the advantage of a much more delicate recovery mechanism than has the muscle. If the nerve is so delicately protected fatigue can not be discovered satisfactorily or studied successfully by the methods used in studying muscular fatigue.

An important element in the neural mechanism that should be considered in connection with the study of neural fatigue is the central part of the nervous arc. If it could be demonstrated that the nerve fiber is not subject to fatigue it would not follow that the formation of bonds does not deteriorate when work is too continuously prolonged. One's ordinary experience teaches him that the time to do some original problem, or to memorize material is when one is fresh and to do routine mental work when one is not at his best. Yet most of the experiments on mental fatigue have been with tasks of a more or less routine character such as multiplication, addition, judgments, reaction time, etc. The investigations so far reported seem to show that the farther one gets from routine material the more likely he is to find some indications of neural fatigue. Memory which is somewhat removed from routine shows the greatest effects of fatigue when memorizing has been continued over long periods.

Oehrn[1] experimented with lists of 12 digits for two hours and nonsense syllables for 1.5 hours. In the latter case the subjects did lose in efficiency as the work progressed. The relations of the number of nonsense syllables learned in the successive 15 minute periods were as 1000, 855, 826, 754, 666 and 751. In criticizing this work Thorndike says: "The meaning of the results obtained in memorizing nonsense-syllables is doubtful; for, as one learns more and more series, there develop interference and confusion. If tib has been connected with (1) pon zet luf (2) nog, biz, ref, and (3) sib, kol, mek, it thereby makes the learning of tib, wif, tek, saw

[1] The investigations here referred to are all reviewed by Thorndike, 'Educational Psychology,' Vol. III.

or tep, lin, tod wak, or deb, nig, ron, puf; and the like harder. The first hour's learning may then reduce later ability by such interference and confusion, as well as by fatigue proper."

In learning lists of 12 digits the scores of the successive 15 minute periods were as; 940, 852, 938, 867, 960, 908, and 1000.

Vogt worked seventy-five minutes daily for twelve days at memorizing series of twelve digits (against additional distractions during the second half-hour). The results for the first and last five minutes of work give an average loss from the beginning to the end of the same day of 7.5 per cent.; and a gain at the beginning of one day over the end of the preceding day, of 9 per cent.

Vogt's results in the case of memorizing nonsense syllables showed a falling off of 27 per cent. in the last five minutes of a seventy-five minute period (of which two fifths was spent in work against distraction) compared with the first five minutes. When the work of the last five minutes of each day is compared with that of the first five minutes of the day following, the drop is 30 per cent. As has been noted, the interference and confusion resulting from the likeness of the different nonsense syllables may have a part in this.

Woodworth who had some preliminary drill in memorizing numbers worked constantly from 3 to 7 P.M. m mor zing sets of numbers. Outside of a slight decrease in efficiency toward the end which was probably due to lighting conditions there was no sign of inability to work.

Tests of the effect of fatigue from school work or other work during the day on ability to memorize have all been negative. Where memory has been looked upon as an index to measure fatigue from other lines of work, little or no loss in efficiency has been discovered. Where memorizing was continued for long periods, and loss in efficiency in memorizing looked for, some effects of fatigue have been discovered; but in these cases the results have been interpreted by Thorndike as confusion and interference due to the use of the same nonsense syllables in different combinations.

In none of this work as far as we have been able to discover

were tests made of the permanency of the associative bonds formed in the fatigue experiments. It is possible that a subject could by increased effort overcome fatigue effects and do work nearly as efficiently after long periods of application as before; but, if retention tests were made, inferiority might be apparent in the bonds formed after long periods of work. This possibility was the hypothesis that led the writer to try the experiment to be reported in this paper. *Can a person adapt himself to long periods of memorizing so that no loss in efficiency will appear? If so, is the material learned at the end of a long period retained as well as that learned at the beginning?* Experiments have already indicated the probable truth of the first question. Our main purpose will be to answer the second.

Statement of Experiment

The purpose of the experiment was to determine whether long continued work in memorizing would result in the bonds formed late in the learning period being less well retained than those learned in the beginning of the learning period. Since it might be possible for an individual to adapt himself in the learning so that he would seemingly learn as much at the end as in the beginning of the period it was necessary to test the subjects for the amount learned at the time of learning as well as to test them for the amount retained after a suitable period of time had elapsed. The relative degree of retention could then be measured by determining the percentage the retained material was of that learned. For instance suppose that an individual learned five units of material in the first part of the sitting and also learned five units in the last part. If after twenty-four hours he was able to reproduce two units of the material learned in the first part and only one unit of that learned in the last part of the learning period one would have obtained evidence of the lack of permanency of bonds formed in the latter part of the sitting.

In order to fulfill the conditions of such an experiment it was necessary to get material of sufficiently small and sufficiently equal units to enable us to keep a careful record of the material memorized. It was necessary to have enough ma-

terial to last for a long period of time. It must be presented uniformly and yet fast enough to cause the subject to exert himself to his utmost in the learning. The material learned at the beginning of the period must be of the same difficulty as that learned at the end so that any difference in learning would be due to the factor of fatigue and not to difference in the difficulty of materials. By proper apparatus for exposure, through careful selection of materials, through the changing of the order in which the material was used with the different subjects and the utmost care in the keeping of constant conditions every attempt was made to eliminate irrelevant factors.

Material

The material used for memorizing was German-English paired associates. A list of such associates compiled by Professor Cattell was procured by the writer. This list contained one thousand pairs. The list was gone over very carefully and all words eliminated which could have any similarity of appearance or sound. After a careful sifting eight hundred and fifty were left and these were used in the experiment. The words were mixed up by having them on slips of paper. They were then selected at random and arranged into lists of ten. The words which happened to be in any list was a matter of chance as well as the order in which the lists were used. To further avoid the possibility that even by chance some lists might be more difficult than others and thus affect the results the order in which the lists were used was changed with the different subjects. This fact, together with the additional fact that seventeen lists are averaged together in the presentation of the results, rules out any possibility of a difference in difficulty in the material accounting for our results. For instance each subject worked with the entire 850 associates at one sitting. The first figure presented in the results indicates the average of the first 170 associates, the second figure the second 170 associates, etc. Furthermore, the 170 associates which were learned first by one subject were learned in another part of the period by another subject.

Apparatus

Since it was necessary to keep the subjects at work and not take any time off to determine how much was learned it was necessary to keep some record of the bonds as they were formed. It was necessary to have an exposure apparatus which would expose the entire 850 paired associates the proper number of times with absolutely no interruption from the beginning of the sitting to the end. Any interruption to change rolls or cards in the apparatus would have given a rest period and would have been as fatal as would any interruption to have the individual undergo a test to determine the amount learned.

In order to meet these rigid conditions a special type of exposure apparatus was designed.[1] The apparatus is repre-

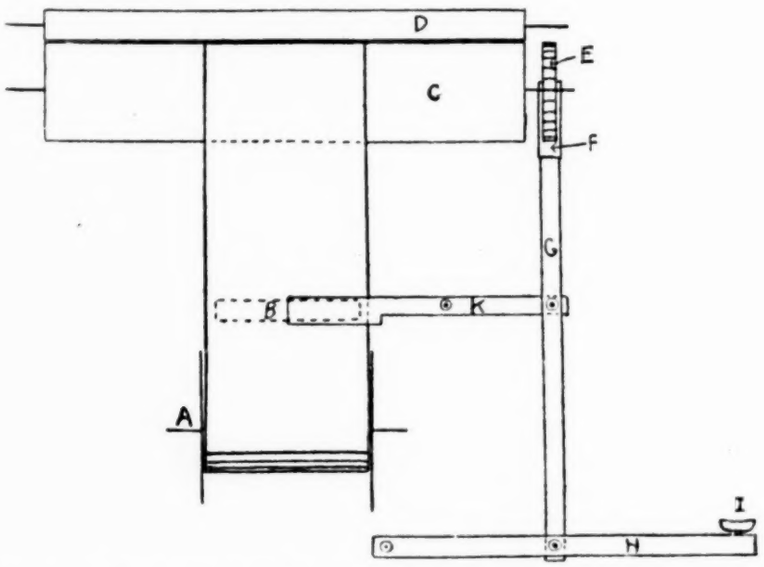


Fig. 1. Paired-associate Exposure Apparatus. Description in text.

sented diagrammatically in Fig. 1. The paired associates were typewritten on a roll of adding machine paper and this roll placed on the bobbin *A*. The end of the roll was then passed behind the exposure opening of the apparatus, which is represented by the dotted lines at *B*, over the larger roller *C* and between roller *C* and the smaller roller *D*. Roller *D* was held tightly against *C* by means of a spring at each end and thus gripped the paper passed between the two. On

[1] The apparatus was constructed in the shops of the Johns Hopkins University.

one end of roller *C* was attached a notched wheel against which played a ratchet, *F*. This ratchet was attached by means of bar *G* to lever *H* so that when pressure was applied to lever *H* at *I* roller *C* turned and pulled the paper the exact distance that a sheet of typewriter paper moves when the double space is used. This movement of the paper changed the material exposed at the apparatus window *B*. Furthermore, it was desired in presenting material to give the subject an opportunity to respond with the proper associate if possible, after the first exposure. This meant that only the first member of the associate be shown until the subject be given an opportunity to respond and then the second member be shown. To accomplish this lever *K* was connected to rod *G* so that when pressure was applied at *I*, half of the exposure opening was covered by *K*. (The apparatus is shown as though *I* were pushed to its lower limit.) When *I* was released the ratchet *F* engaged a new notch in the wheel *E* without turning *C*, and *K* dropped from in front of the opening *B*.

The operation was as follows: The material was all typewritten on the roll of paper in just the order it was to be presented. On the first repetition of a set of associates the experimenter would push the lever *I* and immediately release it. This would bring to the window *B* a complete paired associate—a German word with its English equivalent. After four seconds he would push and again immediately release *I*, thus exposing a second pair. After all ten associates had been shown once they would then appear on the paper in a different order. On the second exposure, however, the experimenter would press *I* down and hold it for two seconds, during which time the subject would try to speak the proper English equivalent of the German word thus exposed. Whether he succeeded or not *I* was released thus showing the English equivalent and two seconds allowed to elapse before it was again pressed and another German word shown. After one set of words had been shown the proper number of times the words "NEW LIST" appeared and immediately a new series was presented. Hence, from the very start to the end of the period no rest was given.

The apparatus was equipped with adjustment mechanisms so as to enable the experimenter during the experiment to make fine adjustments and thus insure proper presentation. The light was artificial and was constant throughout. The correct responses of the subject as the experiment progressed were recorded by the experimenter.

Preliminary Experiments

The preliminary experiments are important in that they indicate conditions under which the hypothesis leading to them is not borne out. They also show the steps which led to the formulation of the experiment in its final form. The results from these experiments when considered in connection with later results give our conclusions a form they would not take were these preliminary results omitted.

Table I

First Two Preliminary Experiments

	Divisions of Learning Period			
	1st Quar.	2d Quar.	3d Quar.	4th Quar.
Av. correct res. in learning per.	7.8	8.0	7.8	7.8
Av. correct res. in relearn per.	4.5	4.0	4.2	4.5
Percentage	57	50	54	58

Subject *A* was used for four different experiments. In the first two she was presented 17 lists of ten associates each. Each list was presented ten times, the order being varied with each repetition. After twenty-four hours, she was given an opportunity to give the English equivalent of each German word. The order in which the material was presented for this recall differed from that in the learning presentation. In each case whether she made the correct response or not she was shown the correct English associate and then later in the same period given a second opportunity for recall. The second experiment was an exact repetition of the first with different material. Table I. divides the material into four groups, excluding the first set of ten associates, in accordance with the order in which they were presented in the learning period. The first line shows the average number of

correct responses made by the subject in learning. The second line gives the average correct responses per list of ten in the recall period. The third line gives the relation of the recall to the learning for each period. It is evident from this table that if fatigue had any effect on retention it was overcome by practice effect, for as a large a percentage was recalled for the last period as for the first.

In the third preliminary experiment the same subject was given a new set of materials in the same way with the exception that only five repetitions were used instead of ten. Two opportunities were again given for recall, the second after a prompting. The results are given in Table II. and indicate that less complete learning did not render the fatigue effect upon retention any more apparent. Sixty-nine per cent. was recalled for the last period as compared with forty-seven for the first period.

TABLE II

THIRD PRELIMINARY EXPERIMENT

	Divisions of Learning Period			
	1st Quar.	2d Quar.	3d Quar.	4th Quar.
Av. correct res. in learning per.	7.7	7.1	6.6	6.9
Av. correct res. in relearn per.	3.6	2.9	4.2	4.8
Percentage.	47	41	64	69

In the fourth preliminary experiment subject *A* was presented 170 nonsense syllable paired associates in the same manner as the German-English associates, ten repetitions were given. After twenty-four hours a recall test was made and not a single associate was recalled. Thirty associates were then selected from the first part of the learning period and thirty from the last part. These sixty associates were mixed up and relearned and the number of errors made in the original learning compared with the errors made in the relearning of the same material. The figures in Table III. indicate that the material learned in the latter part of the first period was not retained as well as that learned in the first part.

TABLE III

FOURTH PRELIMINARY EXPERIMENT

	Errors per List in Learning	Errors per List in Relearning	Percentage
When fresh.....	58	12	21
When fatigued.....	65	20	31

Although this last experiment indicated that the effect sought appeared with nonsense syllables, the work with nonsense syllables was not continued because their use would subject the experiment to the criticism made against them by Thorndike—that they are subject to interference. On the basis of the above experiments it was decided that the subjects probably were not hard pressed enough due to the fact that too many repetitions were given and that probably the work was not long enough continued. It had been first planned to divide the material into sections and try the same subjects with the different sections. Now the material was all arranged so that it could be given to one subject at one sitting and the number of repetitions reduced to four. This meant that the subject would work for nearly four hours and in that time have 850 German-English paired associates presented to him four times. For an individual who had no previous knowledge of German this was no little task.

MAIN EXPERIMENT

Five high-school students were used in this experiment. None of them had any knowledge of German. The learning was conducted on Saturday afternoons beginning at one o'clock. The recall and recognition tests were given the succeeding Monday afternoon. On the first presentation an associate appeared for four seconds and then was followed by another until ten pairs—constituting a list—were presented. Immediately after this first presentation of a list a German word appeared alone for two seconds, and the subject responded with the English equivalent if possible. If he made the correct response the experimenter made the proper record of this fact. Whether he gave the proper

equivalent or not the proper word was then shown by the side of the German word for two seconds at the end of which time another German word appeared alone. This process was continued until all the words in a list had appeared four times, including the first appearance. The only respite between lists was for two seconds when the words "New List" appeared at the opening to warn the subject that a new set of associates was to be shown. This was kept up with absolutely no intermission until the entire 850 associates had been treated in the manner described.

Different orders of presentation were used for the different subjects so as to equalize any chance difference in the difficulty of the materials. The material was not presented to a subject in the relearning period in the same order as in the learning but was mixed—some of that used in the first part of the learning period appearing in the latter part of the relearning.

When the subject reappeared on Monday he was first given a number of German words and told to check the ones he recognized as having been used in the experiment two days previously. This list contained about half the German words actually used in the experiment mixed with the same number of German words that had not been used.

After the recognition test the subject was given an opportunity for recall; that is, an opportunity to give the English equivalent for every German word he had studied two days previously. After a prompting he was given a second opportunity for recall. In the tables the recall scores indicate the number of correct responses given in each list of ten associates on the first presentation. The score obtained upon the second presentation—that is, after one prompting—is termed the relearning score.

Table IV. gives the learning score for the five subjects. *A*, *B*, *C*, *D*, and *E* represent five equal consecutive sections of the learning period. Since 850 associates were used in the experiment each of these consecutive periods is the average score for 17 lists or 170 associates. None of these subjects were given any practice in the work but were started directly

upon the work. Any practice effect is overshadowed by fatigue and there is a gradual diminution in efficiency until the last section of the learning where there is a slight improvement. This indicates that an individual can adapt himself

TABLE IV

LEARNING SCORE

Subject	Consecutive Sections of Learning Period.				
	A	*B*	*C*	*D*	*E*
E	2.10	2.17	1.41	1.49	2.10
F	3.20	3.02	2.02	2.78	3.39
G	3.67	3.94	2.85	3.18	3.25
H	3.55	2.80	2.73	2.06	2.92
I	5.57	4.86	5.96	5.33	5.67
Av.	3.62	3.36	2.99	2.96	3.47

TABLE V

RECOGNITION SCORE

Subject	Consecutive Sections of Learning Period				
	A	*B*	*C*	*D*	*E*
E	5.33	3.33	3.17	3.17	2.34
F	5.25	4.62	4.38	2.88	1.75
G	6.00	4.75	2.50	1.38	0.25
H	4.38	3.38	2.88	3.00	3.00
I	6.36	5.62	5.37	5.37	6.62
Av.	5.47	4.34	3.66	3.16	2.79

TABLE VI

RECALL SCORE

Subject	Consecutive Sections of Learning Period				
	A	*B*	*C*	*D*	*E*
E	0.71	0.00	0.00	0.12	0.23
F	0.41	0.12	0.06	0.30	0.18
G	0.24	0.24	0.12	0.18	0.12
H	0.12	0.06	0.35	0.18	0.18
I	0.53	0.71	0.30	0.30	0.35
Av.	0.40	0.23	0.17	0.22	0.21

fairly well to the conditions of such an experiment and can persist in spite of any fatigue which might be present. Fatigue in this case cannot be of the same sort as physiological fatigue or such adaptation would have been impossible.

TABLE VII

RELEARNING SCORE

Subject	Consecutive Sections of Learning Period				
	A	*B*	*C*	*D*	*E*
E.....	2.37	1.94	1.12	0.88	1.53
F.....	3.12	2.47	2.06	2.00	2.12
G.....	2.59	2.65	1.71	1.77	1.23
H.....	2.59	1.88	1.94	1.82	2.65
I.....	4.94	4.76	4.94	5.12	4.88
Av.....	3.12	2.74	2.35	2.32	2.48

TABLE VIII

RELEARNING SCORES AS PERCENTAGES OF LEARNING SCORES

Subject	Consecutive Sections of Learning Period				
	A	*B*	*C*	*D*	*E*
E.....	113	90	80	59	72
F.....	98	82	102	72	65
G.....	71	67	60	56	38
H.....	73	67	67	88	91
I.....	89	98	83	96	86
Av.....	89	81	78	74	71

Table VI. gives the number of correct associates which the subjects recalled per list of ten. The number recalled is so small that the results are not significant.

Table VII. gives the score after one prompting. This table indicates that the recall after one prompting is slightly less efficient for the material learned toward the end of the learning period. These results are more significant when compared with the learning scores and are expressed as percentages. Table VIII. gives these percentages and shows a progressive deterioration from the beginning to the end of the learning period. It must be remembered that the recall was carried out with the material in a different sequence from that of the learning period. This table indicates that material learned when one is fatigued is not retained as well as that learned when one is fresh. One subject alone presents an exception to the general result. It must be admitted that the deterioration of ability to retain is relatively slight as can be seen from the graphic representation in Fig. 2.

Table V. and the dotted line in Fig. 2 give evidence of a very significant effect of fatigue. The recognition for the German words studied when the subjects were fresh is vastly superior to that for the German words studied when they were fatigued. This is indication that as work progresses and one becomes loath to go on he tends to narrow himself more and more to the specific thing which he is endeavoring

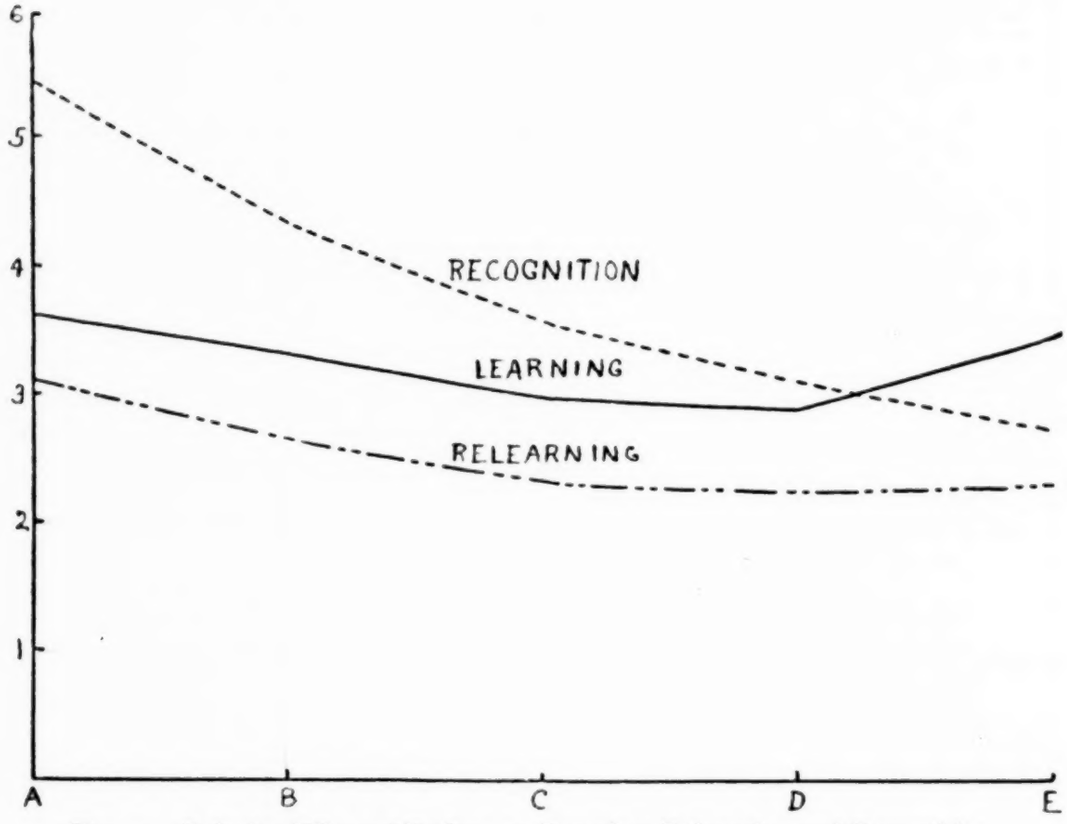


FIG. 2. Relative Effect of Fatigue on Learning, Relearning and Recognition.

to do. When one starts to learn the associates he has some overflow of attention which he can give to details of the elements in the materials with which he is working. As work goes on he excludes more and more of the details which are not strictly in line with the task in hand.

SUMMARY AND CONCLUSIONS

Experiments have shown that one can memorize for long periods of time with little loss in efficiency, but it is quite possible that the permanency of the bonds formed when

one is fresh are more lasting than those formed when one is fatigued. This experiment was designed to ascertain whether material learned at the end of a long period of work is retained as well as that learned at the beginning.

The experiment consisted in having subjects partially learn the English equivalents of 850 German words in one sitting of about four hours. After two days recognition, recall and relearning tests were made to determine the degree of retention for the material learned in the different parts of the learning period.

It was found: 1. There is little loss in efficiency manifested in the learning period. There is a slight drop in the middle followed by a slight spurt toward the end.

2. Of the material learned in the last section of the learning period eighteen per cent. less is retained than that learned in the first section.

3. This result is only apparent when the material is but partially learned, when there is a vast amount of material presented at one sitting and where the work is long continued.

4. Recognition of the material used in the early part of the learning period is much superior to recognition for material used in the latter part. This indicates that adaptation to fatigue involves a narrowing of neural activity to the specific process demanding adaptation.

EFFECTS OF SMOKING ON MENTAL AND MOTOR EFFICIENCY

BY SVEN FROEBERG

Introductory

The question whether or not smoking has any appreciable immediate effect on a person's mental and motor efficiency is one which for several years has been of interest to the writer, both scientifically and personally. This interest was first aroused by certain incidental observations which seemed to indicate a detrimental effect on the higher mental functions.

A series of rough experiments were made on college teachers and students, without, however, obtaining any satisfactory results. Subsequently an opportunity to repeat these on a larger scale and under more favorable conditions was afforded at the psychological laboratory of the University of Michigan. The experiments here reported constitute only a part of the program as originally planned. Change of residence and of work has prevented the completion of the program, and the results are, therefore presented for whatever value they may have for the solution of a problem which just now seems on the point of becoming a social if not a political one.

At the time these experiments were begun there was considerable literature on the subject though little experimental evidence. Such an excellent summary of this literature up to the date of publication was made by Professor Burnham in the *Ped. Seminary* for September, 1917,[1] that it is considered unnecessary to repeat it here.

Such experimental evidence as existed at that time was limited almost entirely to motor functions. The most important was that obtained by Professor Lombard by means of the ergograph. There seemed to be a substantial agree-

[1] Wm. H. Burnham, 'The effect of Tobacco on Mental Efficiency,' *Ped. Sem.*, Sept., 1917, Vol. 34.

ment that the amount of muscular work performed is more or less decreased as a result of smoking.

There is much less agreement with respect to its effect on mental functions. Those who conclude that it is detrimental do so, either because of its detrimental effect on muscular work, or because certain statistical investigations of college students have shown that smokers average somewhat lower in their marks than non-smokers.

It is clear, however, that neither of these reasons is a good one. Mental and motor functions are sufficiently distinct even to a behaviorist, to make it at least conceivable that one function may remain unaffected by conditions which markedly influence the other. And as in many other cases of concomitant phenomena, so also in the case of smoking and low scholarship, it is impossible from the mere figures to determine which is the cause and which the effect. It is quite as conceivable that low scholarship, or the conditions responsible for it, are the cause of smoking as the opposite.

In the absence of satisfactory experimental data the most reasonable conclusions on the subject are perhaps those of Dr. Meylan quoted in the review cited above, to the effect that:

1. The effect of tobacco on adolescents is injurious;
2. On adults the effect is injurious:
   a. In certain individuals suffering from certain nervous affections;
   b. In persons with an idiosyncracy against tobacco;
   c. In all cases where it is used excessively;
3. But that there is no scientific evidence that the moderate use of tobacco by healthy mature men produces any either beneficial or injurious effects that can be measured.

While these experiments were in progress the results of two other investigations in the same field were published. Berry,[1] using himself only as a subject, investigated the effects of smoking on simple addition and found that less time was required and fewer errors made after smoking than when not

[1] Berry, C. S., 'Effects of Smoking on Adding,' *Psy. Bull.*, 1917, Vol. 14, pp. 25–28.

smoking. The fact that every pair of sittings yielded the same general result and that a contrary result was expected by the subject makes these experiments significant as far as this particular subject is concerned.

The other investigation was made by Dr. Bush.[1] He used 17 subjects, 15 who smoked tobacco, 1 who smoked cubebs, and one control, in ten different tests, each test repeated five times after smoking and five times when not smoking. The results show for every test a decrease of efficiency, varying from 3 to 22 per cent., with an average of 10 per cent. for all the subjects taken together. When the subjects are taken separately, as of course they should be, the decrease is from 5 to 18 per cent. when all tests are averaged together, but some of the subjects show no effect, or even a gain after smoking in some of the tests when these are taken separately.

In spite of the apparent care with which this investigation was made there are in it certain defects sufficient to cast serious doubts upon the validity of the results. In the first place the results from the smokers were "corrected" by those from the normal or "control." But the results from only one subject, averaged as they were from only five experiments can scarcely be considered sufficiently "normal" to be used as a standard for the other 15. Secondly, the coefficient of reliability for the individual subjects cannot be calculated, since the variability is not given. Third, the ground upon which certain of the results, which should have been included in the averages, were thrown out as unreliable is not sufficiently explained. To these may be added, the failure to explain how the matter of suggestion was handled, which has proved such a troublesome problem to other investigators.

A third investigation[2] has recently come to the author's attention. The tests were somewhat similar to those used in the present investigation, except that they were fewer and a smaller number of subjects were used. All of the above

[1] Bush, A. D., 'Tobacco Smoking and Mental Efficiency,' *N. Y. Med. Jour.*, 1914, Vol. 99, pp. 519–27.

[2] Effect of Smoking on Mental and Motor Efficiency, *Psychol. Clinic*, June, 1918, and May, 1919. (Oscar J. Johnson.)

criticism apply here also. The superiority of the non-smokers over the smokers amounting to about 1.1 per cent. is too small to be of any significance, and the general conclusion, that "smoking is very detrimental to muscular control and also to the purely mental processes," is not borne out by the figures.

The problem, therefore, still remains to be solved, and the time spent on the experiments reported herewith may not be entirely wasted.

Material and Procedure

The principal difficulty in experiments of this kind, and the one which probably more than any other has served to discourage experiments in this field, is how to eliminate, or at least control, the effect of suggestion. Much of the work on the influence of alcohol and other drugs has been justly criticized for the failure to take this factor into account. Hollingworth in his experiments on the effect of caffeine, and Poffenberger in similar work on strychnine succeeded in eliminating completely the influence of suggestion by the simple expedient of administering the drugs in capsules, and using blank capsules on alternate days. In the case of alcohol similar attempts to disguise the drug have been less successful. Even Dodge and Benedict, who seem to have taken more pains than any previous investigators to neutralize this factor, confess that their mixture of orange juice and other substances succeeded only imperfectly in masking the alcohol. As for tobacco, no attempt whatever seems to have been made in any investigation so far as reported to eliminate or even control suggestion. In one of the investigations mentioned above it is expressly stated, and in the other two it is intimated that a detrimental effect of smoking was expected.

Theoretically there are several methods of eliminating suggestion, but in actual practice serious difficulties were encountered in the case of each one. The first one to suggest itself was to use nicotin-free tobacco on alternate days. It seems that such tobacco has actually been produced, but that there was so little demand for it that it is no longer on

the market. There are several ways of extracting nicotin from the tobacco, but since the flavor and aroma is also changed in the process, this method was not considered advisable. A colleague, evidently himself not a smoker, suggested the use of cigars made of dried cabbage leaves on alternate days, but the suggestion was not adopted. It appears, moreover, doubtful if any of the alleged effects of smoking are due to the nicotin, since ordinarily most, if not all of the nicotin is decomposed in smoking and the small amount which may possibly be volatilized, is again condensed and deposited at the bottom of the pipe or in the unconsumed portion of the cigar. The effects, if any, of smoking are held to be due to the various alcaloid decomposition products of the tobacco, the principal of which is pyridin.

But since the characteristic aroma of the smoke seems to be due not to the nicotin nor any of its decomposition products, but to a substance known as nicotinin or tobacco camphor, it was thought a satisfactory method of control might be found by filtering out the former while leaving the latter in the smoke.

Several methods of doing this were tried, among them that of passing the smoke through a succession of bottles of acidulated water. The method finally adopted consisted in passing the smoke through a glass tube filled with absorbent cotton by which the various alcaloid decomposition products were condensed and absorbed. When smoke thus filtered was passed through a solution of silico-tungstic acid, a very delicate indicator of these substances, no sign of their presence could be obtained.

As finally used the smoking apparatus consisted of a sheet metal tube with a cigar holder attached to a rubber stopper, a long rubber tube filled with absorbent cotton, and a narrow rubber tube to the end of which was attached an ordinary cigarette holder. These were then connected in such a way that the air had to pass through the whole apparatus before reaching the smoker. The positions of the sheet metal tube and the glass tube could be interchanged, thus allowing the smoke to be filtered, when they were in one posi-

tion and allowing it to pass direct to the smoker when in the other. This method was adopted rather than that of dispensing with the filter tube on alternate days, because the difference in "pull" necessary with and without the filter would at once have indicated to the smoker whether the smoke had been filtered or not. There was surprisingly little difference in taste between the smoke thus filtered and ordinary smoke.

In spite of this, however, the method was not entirely satisfactory. The filtered smoke had less visibility than the unfiltered. In some of the subjects this led to more vigorous efforts because they feared the cigar was "going out." In at least one case this was sufficient to indicate to the subject when the smoke was filtered and thus destroyed the value of the apparatus. For this reason only three of the subjects used the apparatus. Such possible methods of overcoming this difficulty as blindfolding the subjects while smoking or smoking in the dark room were not resorted to, since these would have introduced other complicating factors of their own.

In the other subjects counter suggestion was depended upon to eliminate any effects of suggestion. The subject was told that so far there was no experimental evidence to show that smoking had any effect, either favorable or unfavorable, on mental efficiency and were asked to adopt a strictly neutral attitude on the subject. Judging by the statements of the subjects themselves there is every reason to believe that this was also the attitude actually adopted by them. They were, of course, not allowed to see the results while the experiments were in progress.

Cigars rather than cigarettes or pipe were used, because it was desired to eliminate the possible effects of the decomposition products of the paper in the case of the cigarettes, and because a pipe would mean interruption while refilling and relighting. The cigars used were—or at least advertised as—Key West Havana, sold for 5 cents, and varied in weight between 6 and 7 grams. Each cigar as well as the unsmoked part was carefully weighed, and the amount of tobacco actu-

ally smoked in each experiment thus determined. The average amount consumed was between 4.5 and 5.5 grams. The nicotin content was analyzed by one of the subjects, who was also taking courses in organic chemistry, and was found to be between 2.29 and 2.33 per cent. dry weight. The smoking period was approximately 30 minutes.

The tests were selected so as to cover as wide a range of mental and motor functions as possible within the 30 minutes alloted to the tests. All were selected from either Whipple's Manual of Mental and Physical Tests or Woodworth and Well's Association Tests, with the exception of the test on attention and the completion test. The motor tests comprise a test of steadiness of motor control, a test of accuracy of voluntary movement, and a test of speed of voluntary movement, listed as Nos. 13, 12 and 10 respectively in Whipple's Manual.

The remaining tests have been classed as Mental. Of these one was intended to be a test of attention and the rest tests of association. The attention test was devised by Prof. J. F. Shepard and used by Dockery in his work on fatigue, but because of imperfections in the apparatus the results of this test have not been included.

The memory test consisted in reading aloud, at the rate of one per ¾ second, ten consonants arranged in such a manner that all combinations likely to suggest meanings were excluded. Two different series were used in each test, and Spearman's foot rule was used in scoring the results.

The association tests comprise, the uncontrolled association, opposites, addition and completion tests. The uncontrolled association test is the same as No. 33 of Whipple's Manual, except for the fact that the stimulus words were taken from the Kent-Rosanoff list.

The addition test was essentially the same as the constant increment test of Woodworth and Wells, except that only the numbers from 11 to 82 not ending in zero were used. The constant increment was 17.

The opposites tests were those of Woodworth and Wells.

The completion test was added because it was desired

to include a test of sufficient difficulty to require the use of the most complex functions of intelligence, and because the results of the Trabue Language Scales have shown that this test correlates more highly with general intelligence than any other single test. The test consisted in supplying the missing words in a paragraph of 100 words from which about 15 had been eliminated. The eliminations had been made in such a way that the completion could usually be made by any one of several words but with varying degrees of appropriateness. The subjects were instructed to have regard for both speed and accuracy. The word supplied was graded on a percentage basis by 5 separate judges, except in one set, where one judge made two gradings with an interval of a month between. To combine speed and accuracy into a single score it was assumed that, other things being equal, these vary inversely and in proportion to their coefficients of variability. Both times and grades were then expressed in terms of the Woodworth reduced measures,[1] *i.e.*, the deviation from the average divided by the variability. In the first the standard deviation was used as the coefficient of variability, but since the result was found to be practically the same when the average deviation was used, the latter was adopted in the second set, thereby saving a great deal of time and labor. The measures thus obtained the times were substracted from the corresponding measures for the grades, and the result corrected for unequal difficulty by substracting the similar scores of five control subjects. The reliability of the scores is of course affected not only by the variability of the subjects themselves, but also by the variability of the control subjects and the judges. To make results of this test comparable in reliability with that of the others it would be necessary to increase the number of judges and control subjects as well as the number of test subjects. But as the labor of preparing and scoring the results of this test was already more than twice that of all the test combined, it was not thought that the slight gain in reliability thus obtained would repay the additional work.

[1] R. S. Woodworth, 'Combining the Results of Several Tests,' etc., *Psychol. Review*, Vol. 19, pp. 96 ff.

Since it was necessary to use the same material repeatedly in the memory, opposites, and addition tests, the effect of practice had to be guarded against. To this end the order of the consonants, words, and numbers was different in each experiment.

The 30 minutes preceding the tests were spent in smoking, or in light reading such as newspapers and magazines or in conversation on non-smoking days. The tests usually lasted 3 minutes.

The subjects were all students in the elementary psychology course at the University of Michigan, varying in age from 19 to 24. All of them were what may be termed occasional smokers and had consequently all acquired a tolerance to tobacco, but none were habitual smokers at the time. The original plan provided also for experiments on the latter kind. A certain amount of scholastic credit was given to these subjects for taking part in the experiments, and for this reason they were required to sign an agreement to abstain from smoking or indulging in any alcoholic beverages, tea, or coffee previous to the experiments on the days these were scheduled, to adopt, as far as possible, regularity in the hours or work, sleep, etc., and to observe certain other conditions necessary for the success of the investigation. No experiments were made on days when these conditions did not seem to have been fulfilled.

In order to allow for any possible influence of atmospheric conditions, a complete record was also kept of temperature, pressure, humidity and light conditions.

The pulse rate of each subject before and after smoking was observed at least once during the experiments. In every case except one it was found to be higher after smoking than before.

Results

The results of the first five subjects are given in Table I., the first column of figures being the average record of 8 normal days, the second the difference between this and the score on smoking days, and the third the P.E. of this difference. The sign before the figures in the second column of

TABLE I

Subjects	Steadiness			Coördination			Tapping			Memory			Free Association			Addition			Opposites			Completion	
	Av.	Diff.	P.E.	Av.	Diff.	P.E.	Av.	Diff.	P.E.	Av.	Diff.	P.E.	Av.	Diff.	P.E.	Av.	Diff.	P.E.	Av.	Diff.	P.E.	Diff. Dev./σ	P.E.
Be.....	34.0	+10.5	7.8	3.65	−1.08	.27	325.5	− .8	4.0	61.4	− 4.7	3.3	39.4	− .1	.95	74.9	+ .4	3.9	27.9	− .3	1.3	−1.30	.36
Bl.....	64.0	+74.5	7.6	6.63	−2.49	.26	299.6	+ 3.1	3.5	46.9	+ 1.2	2.5	53.5	−1.1	3.4	73.9	−4.6	5.3	40.9	−2.6	1.9	− .11	.38
La.....	8.0	− 1.6	2.5	5.84	− .33	.24	300.9	+14.1	5.4	59.8	+ 2.9	3.1	25.8	−2.5	1.3	166.5	−5.0	5.8	34.3	− .4	1.6	+ .74	.36
Ly.....	15.4	+23.6	3.2	4.35	−1.06	.28	316.0	− .5	3.3	77.6	− 2.8	3.2	37.1	−1.0	1.8	107.0	+3.9	7.5	32.1	− .5	1.3	− .04	.42
M.....	31.2	+16.8	6.1	4.57	− .92	.30	301.4	+ 9.2	6.2	74.4	−10.7	4.4	45.0	−3.2	1.7	81.9	+2.5	5.6	32.4	+4.0	1.4	− .30	.40
Av.....	30.6	+36.8	5.3	5.01	−1.18	.27	308.7	+ 5.0	4.5	64.0	− 2.8	3.3	40.2	−1.4	1.9	100.8	−1.1	5.6	33.5	.0	1.5	− .20	.38

figures indicates whether the score after smoking is higher or lower than the normal.

The score in the steadiness test indicates the number of contacts per minute, in the coördination test the distance at which the first contact was made, in the tapping test the number of taps made during the period, in the memory test the grade according to the Spearman foot rule formula, in the free association test the numbers of words during the ten-second period, in the addition test the time in seconds, in the opposites test also the time in seconds, and in the completion test the combined efficiency score as explained above.

It will be seen from the table that in steadiness all except Be and La and in coördination all except La show a marked decrease in efficiency after smoking. In this respect the results agree with those of Professor Lombard referred to above as well as with a widespread belief among marksmen and training camp instructors that smoking is detrimental to delicate muscular control.

Assuming that a difference in order to be significant should be at least 3 times the P.E. it appears that smoking has no appreciable effect on the speed of movement, at least the sort of movement made in tapping.

The results of the purely mental tests agree neither with those of Bush and Johnson on the one hand nor with those of Berry on the other. In some cases the score is higher after smoking, in other cases lower. But in no case, except that of Be in the completion test, is the difference as much as 3 times the P.E.

It was then suggested that this difference in the results in the two kinds of tests might possibly be due to the fact that the effects of smoking had disappeared by the time the mental tests were given, since these came last on the program. In order to remove all uncertainty on this point, the experiments were repeated on a new set of subjects, using only the mental tests. This made it possible to give the tests twice each day, immediately before and immediately after smoking. Each series of tests occupied 15 minutes and the

half-hour intervals between them was spent in smoking or, on alternate days, in conversation or light reading. The method adopted in scoring was that used by Dodge and Benedict in their investigation of the effects of alcohol, viz. $E = (S - Ns) - (N_2 - N_1)$, where E represents general efficiency, S the results after smoking, Ns the results before smoking, N_1 the results of the first series on non-smoking days, and N_2 those of the second on non-smoking days. Twenty-four series in all were given, yielding an average from six differences between smoking and non-smoking days. In all other respects the conditions were the same as before. The results are given in table II.

TABLE II

Subject	Memory		Free Association		Addition		Opposites		Completion	
	Av. Diff.	P. E.	Av. Diff.	P. E.	Av. Diff.	P. E.	Av. Diff.	P. E.	Diff.	P. E.
Bu	59.6 − 14.0	6.1	40.3 − 6.3	2.0	62.7 + 2.6	4.4	17.3 − .6	.6	−1.51	2.54
H	69.8 − 14.0	6.3	31.7 − 3.0	1.1	49.9 − 1.8	1.0	16.8 + 1.9	.4	−.92	1.20
P	64.8 − 24.0	3.8	44.8 − 7.8	2.8	53.9 + 4.2	2.3	20.0 + 1.2	1.2	− .37	1.33
R	51.6 − 13.2	5.1	23.0 − 3.5	1.8	44.8 + 1.8	2.2	13.1 + .3	.8	− .12	3.31
W	50.7 + 16.8	6.4	23.0 − 1.8	.8	34.9 + 6.1	3.8	11.9 + .6	.4	−2.29	1.74
Av.	59.3 − 9.7	5.5	37.6 − 4.5	1.7	49.2 + 2.6	2.7	15.8 + .7	.7	−1.04	2.02

The results differ in no respect from those of the first set. If the two sets of mental tests are taken together it is seen that in only 4 cases out of 50 does the difference exceed 3 times the P.E., in 26 cases it does not exceed the amount of the P.E., and in 12 cases the result gives a higher score after smoking.

There is, on the other hand, no clearly indicated improvement such as the 7.7 per cent. found by Berry, the nearest approach to it being subject La, whose differences, however, in no case meet the requirements of reliability mentioned above.

Two subjects, P and R, show throughout a loss after smoking, but in only one case is this loss greater than three times the P.E. It is possible that if more practiced subjects had been available and the repetitions of the tests increased, the chance variability might have been reduced sufficiently to show significant differences.

It is also conceivable that in such a case individual differences among the subjects might be discovered such as that suggested by subject La on the one hand and subjects P and R on the other. It might be expected that some habitual smokers would show improvement after smoking and that the opposite would be true of non-smokers. Experiments on such subjects were contemplated in the original plan.

As far as the results of these experiments go, and as far as the immediate effect of smoking is concerned, they tend to support the conclusions of Dr. Meylan quoted above: that, except when used in excess, by adolescents, by persons having an individual idiosyncrasy against tobacco, or by persons suffering from certain nervous affections, there is no scientific evidence that the moderate use of tobacco, in smoking produces any either beneficial or injurious mental effect sufficiently great to be measured.

A STUDY OF OCULAR FUNCTIONS WITH SPECIAL REFERENCE TO THE LOOKOUT AND SIGNAL SERVICE OF THE NAVY

BY C. E. FERREE, G. RAND AND D. BUCKLEY

Bryn Mawr College

The incentive for this work was the need for establishing a system of testing for those branches of service in the Navy requiring especially keen scotopic or low illumination acuity. The first step towards the accomplishment of this purpose was the devising of a suitable apparatus and test method. The request for an apparatus came to us from the head of the Eye Division of the United States Naval Hospital at Washington. The apparatus was described in a former paper, "An Apparatus for Determining Acuity at Low Illuminations, etc.," this JOURNAL, 1920, III., pp. 59–71. A further need was to find out what range of difference in scotopic acuity might be expected among eyes graded as fit on the basis of the tests of other functions and capacities. A consideration of this need has led us to make a preliminary survey of eyes graded as normal with regard to photopic acuity and other commonly tested functions in order to determine whether such eyes may be expected to show a significant difference in keenness of functioning at low illuminations.

In a thorough test for vocations requiring keenness of discriminations at low illuminations, the following points should be taken into account: (1) the minimum amount of light required to discriminate the test-object before adaptation; (2) the minimum amount after a properly selected period of adaptation; and (3) the rapidity as well as the amount of gain in acuity in the process of adapting. Determinations covering all of these points have been made in this study.

The Range of Illumination Required by Normal Eyes for the Discrimination of the Standard Test Object

In making these determinations three test-objects were used: the Snellen chart and two single test-objects which could be rotated into different positions—the letter E and the international test-object, the broken circle, each subtending a 5-min. angle. In case of the latter two, the task required of the observer was to indicate roughly the direction in which the opening in the test character was turned, an objective check being had on the correctness of the judgment. The determinations were made at the beginning and end of a 45-min. adaptation period. It is obvious that the results at low illumination should be influenced by the refraction condition of the eye as well as by its light sensitivity and the individual differences in the effect of light sensitivity on acuity. In order that the observers could be chosen so that the results would represent the differences which may occur among eyes having normal or better than normal photopic acuity, each eye was refracted and the acuity was taken under 5 foot-candles (53.8 meter-candles) of light. In the first series of tests 22 observers were used ranging from 18 to 28 years of age. Results were obtained for both eyes and for each eye separately. Of the eyes used, 75.7 per cent. would be rated in the Snellen scale as having 6/4 acuity; 13.5 per cent. as having 6/5 acuity; and 10.8 per cent. as having 6/6 acuity. It was our intention to use throughout only eyes which could be ranked as Grade A with regard to photopic acuity.

The results of these determinations show a greater range of individual difference for the broken circle than for the Snellen chart or the letter E (905 per cent. for the broken circle, 548 per cent. for the letter E and 357 per cent. for the Snellen chart). This superior showing for the broken circle is perhaps in accord with the general finding that the broken circle as a test-object picks up smaller differences in acuity than either of the other two test-objects employed. These differences too, it will be remembered, are amplified in the

present case by the fact that they are read on the illumination scale—an amplifying scale—and not on a scale which sustains a 1 : 1 relation to acuity. Inasmuch as the greater sensitivity was shown in these preliminary experiments by the broken circle as a test-object, 15 additional observers (photopic acuity, 6/4) were employed using this test-object alone.

Space will be taken here only for a brief general statement of the results for this latter series of determinations alone. (1) The individual differences in the minimum illumination required for the discrimination of the test-object before the period of dark adaptation fell between 0.70 and 5.29 meter-candles, a range of 657 per cent.; after the period of dark adaptation it fell between 0.32 and 2.2 meter-candles, a range of 593 per cent. A greater range of individual difference, it will be noted, was found for the tests taken before the period of dark adaptation than for those taken after the 45-minute adaptation period. This was doubtless in part due to the lack of careful standardization of the initial sensitivity by a period of preadaptation to light of a fixed intensity and to small individual differences in photopic acuity revealed by the more sensitive method of testing; and in part to individual differences in the amount and rate of adaptation. A careful initial standardization of sensitivity was purposely avoided in this preliminary work with the apparatus in order more closely to approximate the rough conditions of testing which are apt to prevail in the selection of men with reference to vocational fitness. The results of these determinations are shown in Table I.

(2) Thus far without exception the two eyes of the same observer have required a different amount of light just to discriminate the test-object. This difference has ranged after adaptation from 19 to 54 per cent. of the amount of light required for the better eye.

(3) A question is often raised with reference to points of advantage of binocular as compared with monocular seeing. In 6 per cent. of the number of cases tested, the binocular result after adaptation was equal to or approximated the result for the poorer eye; in 88 per cent. of the cases it was better

TABLE I

SHOWING THE AMOUNT OF LIGHT REQUIRED JUST TO DISCRIMINATE THE TEST-OBJECT AT THE BEGINNING OF DARK ADAPTATION AND AT THE END OF 15, 30 AND 45 MINUTES (15 OBSERVERS)

In these experiments there was no standardization of the initial sensitivity by a previous adaptation to an illumination of constant intensity.

Observer	Photopic Acuity	Illumination in Meter-candles Required Just to Discriminate the Test-Object				Difference in Illumination Required at Beginning and at End of 45 Minutes	
		Beginning	At End of 15 Min.	At End of 30 Min.	At End of 45 Min.	Meter-candles	Per Cent.
G.....	6/4	0.70	0.55	0.35	0.32	0.38	118.8
M.....	6/4	1.00	1.00	0.82	0.82	0.18	21.9
Mc.....	6/4	1.24	1.00	1.00	1.00	0.24	24.0
R.....	6/4	1.36	0.60	0.50	0.60	0.76	126.7
L.....	6/4	1.55	1.00	0.88	0.88	0.67	76.1
S.....	6/4	1.75	1.42	0.88	0.88	0.87	98.9
Th.....	6/4	2.11	0.82	0.94	0.94	1.17	124.5
Y.....	6/4	2.11	1.55	1.49	1.42	0.69	48.6
St.....	6/4	2.40	0.60	0.60	0.60	1.80	300.0
Sw.....	6/4	3.43	2.17	2.20	2.20	1.23	55.9
K.....	6/4	3.90	2.81	2.40	2.10	1.80	85.7
T.....	6/4	3.97	1.18	0.82	0.88	3.09	351.1
Sm.....	6/4	4.10	3.80	1.40	1.30	2.80	215.4
W.....	6/4	4.20	1.40	0.76	0.76	3.44	452.6
Ba.....	6/4	5.29	2.11	2.02	2.11	3.18	150.7

than the result for either eye; and in the remaining 6 per cent. of the cases it was intermediate to the results obtained with the two eyes separately. In none of the cases tested separately was it equal or approximately equal to the result for the better eye. In the 88 per cent. of cases referred to, less light was required for the discrimination of the test-object by the two eyes than by the better eye alone by amounts ranging from 14.5 to 67.3 per cent.

In order to get a rough idea of the grouping of the 15 observers with reference to the minimum amount of light required to meet the standard of acuity imposed by the test, before and after the period of dark adaptation, they have in each case been divided into six equally spaced groups, each group for the work before adaptation covering a range of 1 meter-candle and for the work after adaptation a range of 0.4 meter-candle. For the tests before adaptation 13.3 per cent. fall in the first or best group; 26.7 per cent. in the second group; 20 per cent. in the third group; 20 per cent.

in the fourth group; 13.3 per cent. in the fifth group; and 6.7 per cent. in the sixth group. For the tests after adaptation 6.7 per cent. fall in the first group; 20 per cent. in the second group; 40 per cent. in the third group; 13.3 per cent. in the fourth group; none in the fifth group; and 20 per cent. in the sixth group. A graphic representation of this grouping is shown in Fig. 1.

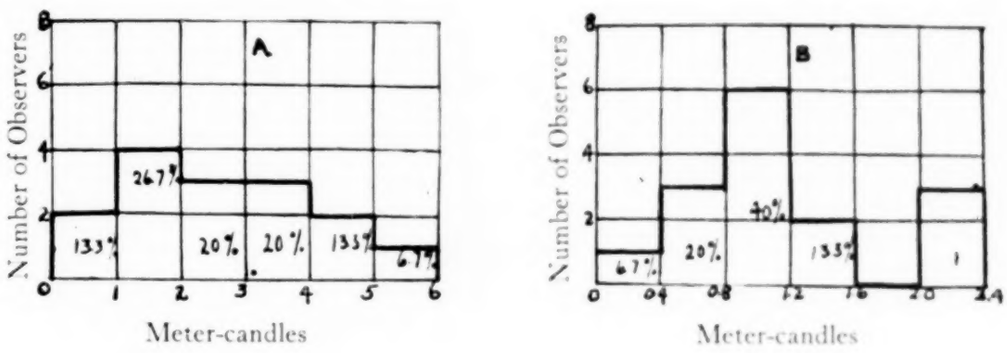


FIG. 1. Representing the relative distribution of 15 observers graded with reference to the minimum illumination required just to discriminate the test-object: *A*, before adaptation; *B*, after a 45-minute period of dark adaptation.

INDIVIDUAL VARIATIONS IN THE AMOUNT AND RATE OF ADAPTATION IN TERMS OF EFFECT ON ACUITY

The preceding experiments furnish data with regard to the minimum amounts of light required just to discriminate the different test-objects at the beginning and end of the 45-minute adaptation period. In case of the 15 observers tested with the broken circle in the final series of experiments at the beginning of dark adaptation and at the end of 15, 30 and 45 minutes, the minimum difference in this amount of light was 0.18 meter-candle or 22 per cent. of the amount required at the end of the adaptation period; the maximum difference was 3.44 meter-candles or 452.6 per cent. of the amount required at the end of the adaptation period. These results are shown in Table I.

In addition to these experiments a special adaptation series was run in which the minimum illumination required just to discriminate the test-object was determined at the

beginning of the adaptation period and at the end of 5, 10, 15, 25, 35 and 45 minutes. In order to standardize the initial sensitivity of the eyes of the observer, a preadaptation period of 20 minutes was given to 80 foot-candles of light (vertical component), the skylight illumination of an optic's room on a medium bright day. A few of the results obtained are represented in Table II. and Fig. 2. These results, it will be remembered, represent adaptation only as it affects acuity, which is the effect of greatest importance to the special work for which the apparatus was devised and the effect with which we are the most concerned for the greater part of our working lives. A comparison of these results with those of similar

TABLE II

SHOWING THE AMOUNT OF LIGHT REQUIRED JUST TO DISCRIMINATE THE TEST-OBJECT AT THE BEGINNING OF DARK ADAPTATION AND AT THE END OF 5, 10, 15, 25, 35, AND 45 MINUTES (6 OBSERVERS)

In these experiments the initial sensitivity was standardized by 20 minutes preadaptation to 80 foot-candles of light (vertical component), the illumination of a skylight optics room on a medium bright day.

Observer	Photopic Acuity	Illumination in Meter-candles Required Just to Discriminate the Test-object							Difference in Illumination Required at Beginning and at End of 45 Minutes	
		Beginning	At End of 5 Min.	At End of 10 Min.	At End of 15 Min.	At End of 25 Min.	At End of 35 Min.	At End of 45 Min.	Meter-candles	Per Cent.
I . . .	6/4	0.55	0.505	0.42	0.35	0.32	0.35	0.35	0.20	57.1
II . . .	6/4	0.705	0.42	0.42	0.32	0.32	0.32	0.38	0.325	85.5
III . . .	6/4	1.06	0.76	0.60	0.46	0.35	0.46	0.46	0.60	130.4
IV . . .	6/4	1.12	0.82	0.52	0.32	0.32	0.38	0.38	0.74	194.7
V . . .	6/4	1.62	1.12	0.94	0.55	0.60	0.60	0.55	1.07	194.5
VI . . .	6/4	2.20	1.12	1.14	1.18	1.36	1.24	1.24	0.96	77.4

series in which the object is to measure the increase in light sensitivity as a result of dark adaptation shows that just as acuity increases slowly with increase of illumination (except at very low illuminations) so also does it increase slowly with increase of sensitivity to light. That is, the eye does not gain in acuity by adaptation nearly so fast as it gains in light sensitivity.

In Fig. 2 the actual amounts of illumination required just to discriminate the test-object are plotted against time of

adaptation. It thus affords a comparison of the position of the minima of the different observers in the illumination scale and comprehends data from which the following points can be determined: (*a*) their relative ranking with regard to

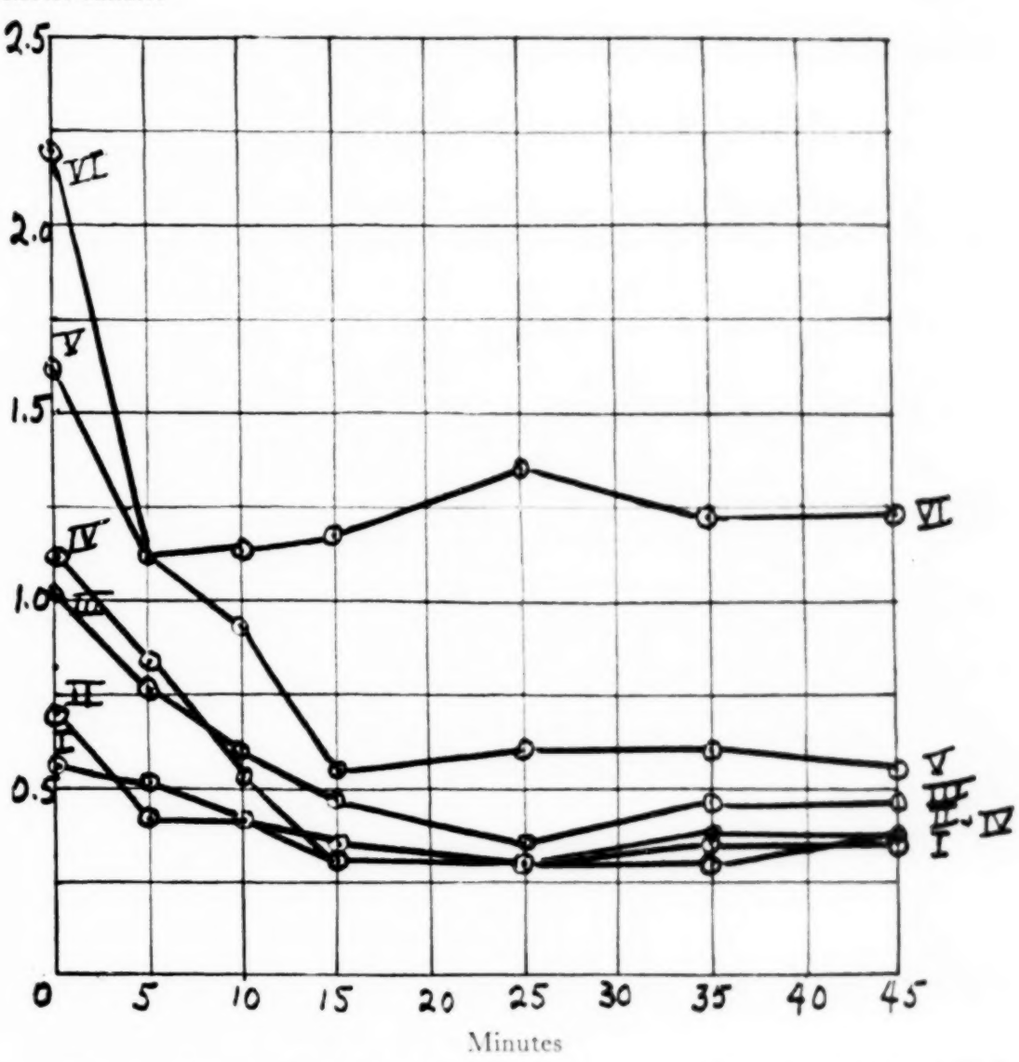


FIG. 2. Curves showing the decrease in the amount of light required just to discriminate the test-object as the result of dark adaptation. In order to standardize the initial sensitivity, the eye was preadapted in each case for 20 minutes to 80 foot-candles of illumination (vertical component).

scotopic acuity before and after adaptation, rated on the illumination scale; (*b*) their light sensitivity before and after adaptation insofar as it affects the minimum amounts of light required to discriminate the test object; and (*c*) their relative amounts of change in sensitivity, measured in terms of effect

on acuity, read on the illumination scale, due to adaptation. All of these features are important for vocational and clinic classification. In order to make these results more directly comparable with reference to the last of these points, namely

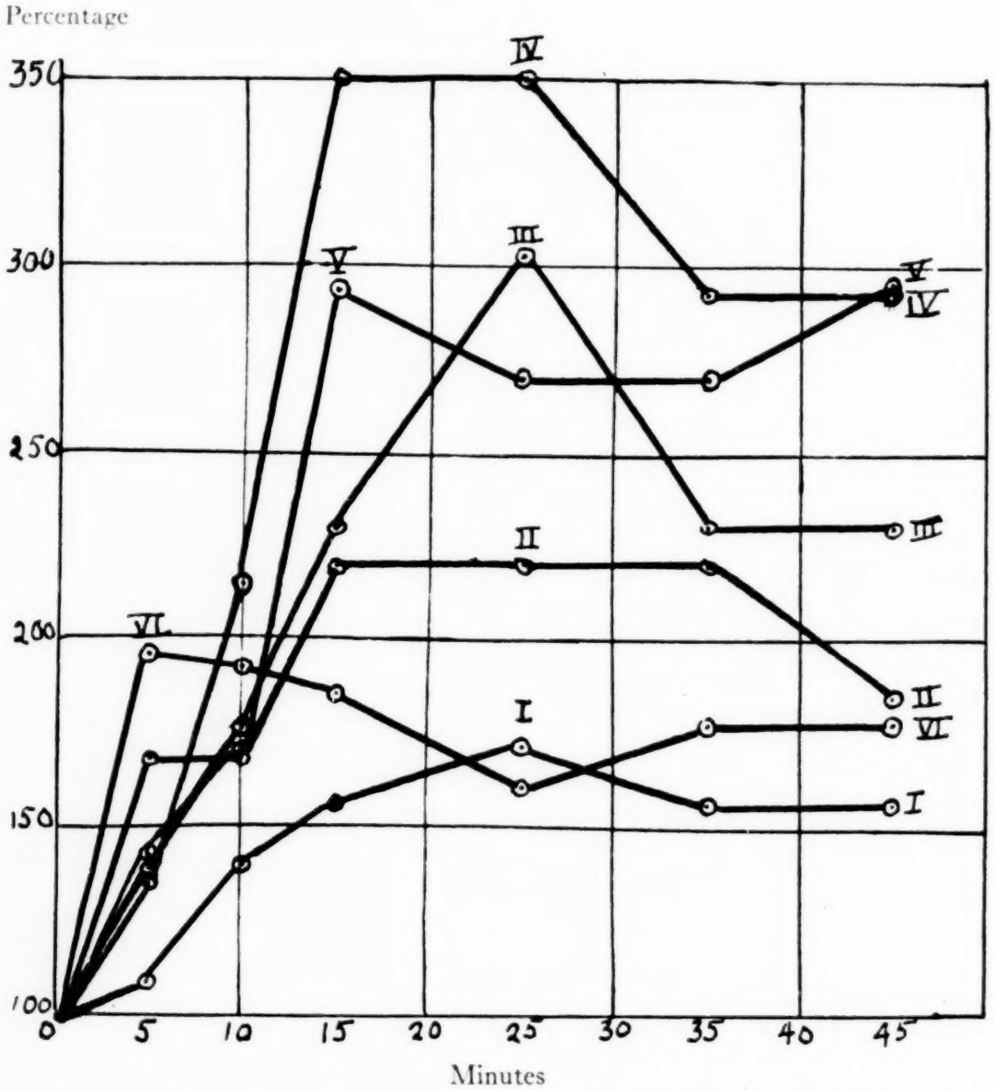


FIG. 3. Curves sho ing the increase in sensitivity as the result of dark adaptation, the reciprocal of the ount of light required just to discriminate the test-object being taken as the measure of sensitivity. Percentage increase in sensitivity is plotted against time of adaptation. The initial sensitivity of the eye was standardized in each case by 20 minutes of preadaptation to 80 foot-candles of light, vertical component.

relative amounts of change in sensitivity due to adaptation, the ratios or percentages of increase in sensitivity are plotted in Fig. 3, the reciprocals of the minimum amounts of light required just to discriminate the test-object being taken as

the measure of sensitivity. That is, the ratio or percentage change in the value of these reciprocals is plotted against time of adaptation, the curves beginning at a common point or unit ratio. The relative ratings with regard to the second point could of course be represented by plotting the reciprocals themselves. Space will not be taken here for this representation.

It will be noted that the greater part of these observers reach their maximum acuity at the end of 15 minutes of adaptation and that some even show a lower acuity if the series is continued beyond this time. The loss in the latter case is doubtless due to fatigue of the muscles of adjustment. That is, in case of the observers more susceptible to muscle fatigue, the loss of acuity due to fatigue more than compensated for the small gain in light sensitivity after the first 15–25 minutes. In this connection it may be noted that the muscle strain imposed by taking the acuity at the minimum illumination is much greater than at the illuminations ordinarily used. Even with a 5–10 minute rest period between determinations and a 2-second interval between the individual observations making up one determination, a very noticeable fatigue was present at the end of the 45-minute series.

In testing fitness for the lookout and signal service work of the Navy, night flying, and for other work and vocations that require the keen discrimination of objects when small amounts of light enter the eye, it was deemed better to make the tests in terms of acuity rather than of the light sense. Retinal sensitivity is only one of the factors in the eye's power to see its objects at low illuminations. For example, we frequently find observers with an excellent light sense whose scotopic acuity or power to discriminate objects at low illuminations is poor. Indeed, as shown in this and former papers, slight differences in resolving power correspond to relatively large differences either in illumination (except for very low illuminations) or retinal sensitivity in their effect on the eye's power to see clearly at low illuminations. Any test therefore for fitness for tasks or work requiring the power to see clearly at low illuminations is far from complete which

leaves out of account resolving power and the varying effect of changes in illumination or in light sensitivity on acuity. The acuity test, on the other hand, includes all of the factors involved in seeing and in the exact proportions in which they are contributory to seeing. Moreover, it is much better directly to determine the candidate's power to see clearly at low illuminations than to try to infer this from a light sense test and data on acuity taken at higher illuminations. This was, we may say, also the point of view of the Naval authorities under the auspices of whom we undertook to devise an apparatus suitable for testing acuity at low illuminations.

THE BLOCK-DESIGN TESTS

BY S. C. KOHS[1]

Portland, Oregon

A brief presentation of the Block-design Tests will be attempted in this article. These tests fall in the category of 'performance tests' and have been standardized to measure intelligence. They have been purposely devised to eliminate the factor of language. In this attempt the writer believes he has been especially successful since the instructions themselves may be given entirely through pantomime and imitation.

There has indeed been, and there still is, a great need for tests such as are here presented. In the longer monograph which the writer is preparing for publication, there will be more detailed treatment of many topics, such as the definition of intelligence; an analytic criticism of current methods of standardization; suggested newer statistical procedure; the relation between language ability, performance and intelligence; and other pertinent material.

The content of the present article has been divided into six sections:

(*A*) The Test Material:
  1. The Blocks.
  2. The Designs.

(*B*) The Directions for Applying the Tests:
  1. For Subjects Who Can Understand Spoken Language.
  2. For Subjects Who Do Not Know the Names of the Colors.
  3. For Subjects Who Cannot Understand Spoken Language.

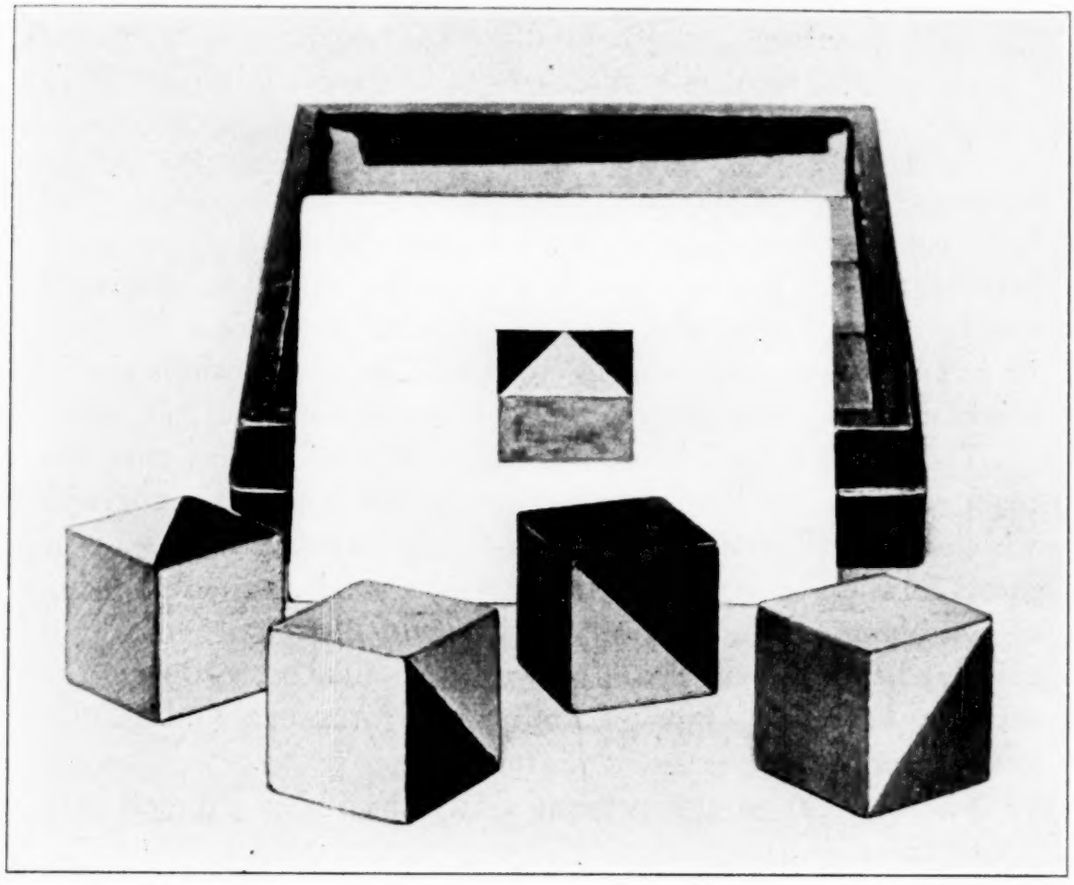
(*C*) The Score Card and Methods of Scoring.

(*D*) The Norms.

[1] Psychologist to the Court of Domestic Relations.

(*E*) The Reliability of the Tests.
(*F*) Serviceability.

In the promised monograph more complete details will be presented which would be out of place in this brief article.



The Block-Design Test.

(*A*) The Test Material

1. *The Blocks*

The Blocks which are used are manufactured by the Embossing Co., and may be secured at any of the large department stores and at various distributing centers of Milton Bradley's. There are sixteen cubes of one inch dimension and all are painted as follows:

One side red
One side blue
One side white
One side yellow

One side blue and yellow (divided on the diagonal)
One side red and white (divided diagonally)

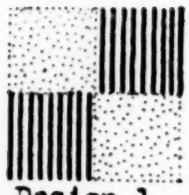
The character of the colors is indicated on the page of designs (pp. 360–1) in this article. A slight difficulty experienced by possibly one or two subjects out of every 100 was a just perceptible but nevertheless disconcerting difference in shade between the blue and yellow on the full faces and the same colors on the diagonal sides. This can be remedied in the later standardization of the test material. One set of the blocks will last through the examination of from four to five hundred children without showing much wear and tear. After that the cubes can be repainted without difficulty.

It is interesting to watch the response of children and even adults when they are given colored cubes to handle. There is no doubt that an appeal exists which touches the roots of some very fundamental original tendencies. Of all the subjects tested, not one has manifested any absence of a desire to *combine* these cubes in some fashion. The experimenter needs only to direct this natural interest toward a specific end and then apply a scientific measuring technique to evaluate the results.

2. *The Designs*

In Chart I. the seventeen designs utilized in this test are represented. The Arabic numerals designate the final numbering of each design. The original number was 35 but fifteen were eliminated in a few of the early preliminary testings. The designs are graded in difficulty which increase by modifying the designs at various stages in the following manner:

1. By the use of the full colors;
2. By the use of few diagonaled sides;
3. By the use of all diagonaled sides;
4. By turning the design on one of its corners;
5. By eliminating the outside boundary line;
6. By increasing the number of blocks to be used;
7. By increasing dissymmetry in design;
8. By decreasing the number of different colors used in each design.

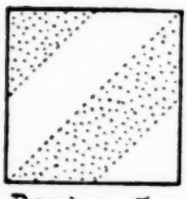


Design 1

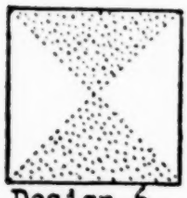
Design 2

Design 3

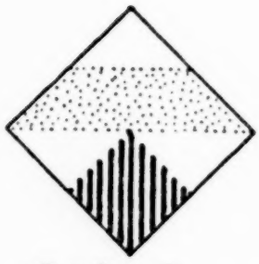
Design 4



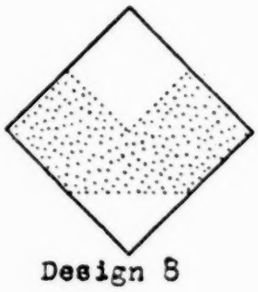
Design 5



Design 6



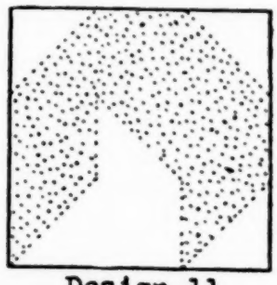
Design 7



Design 8

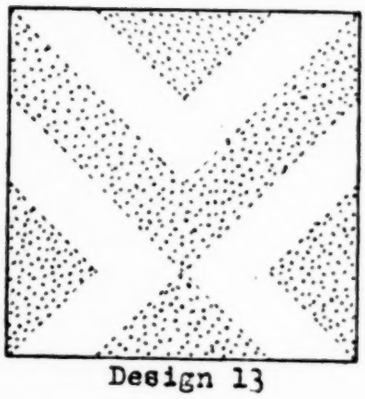
Design 9

Design 10

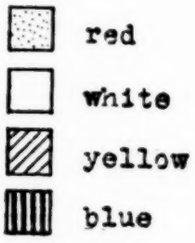


Design 11

Design 12



Design 13



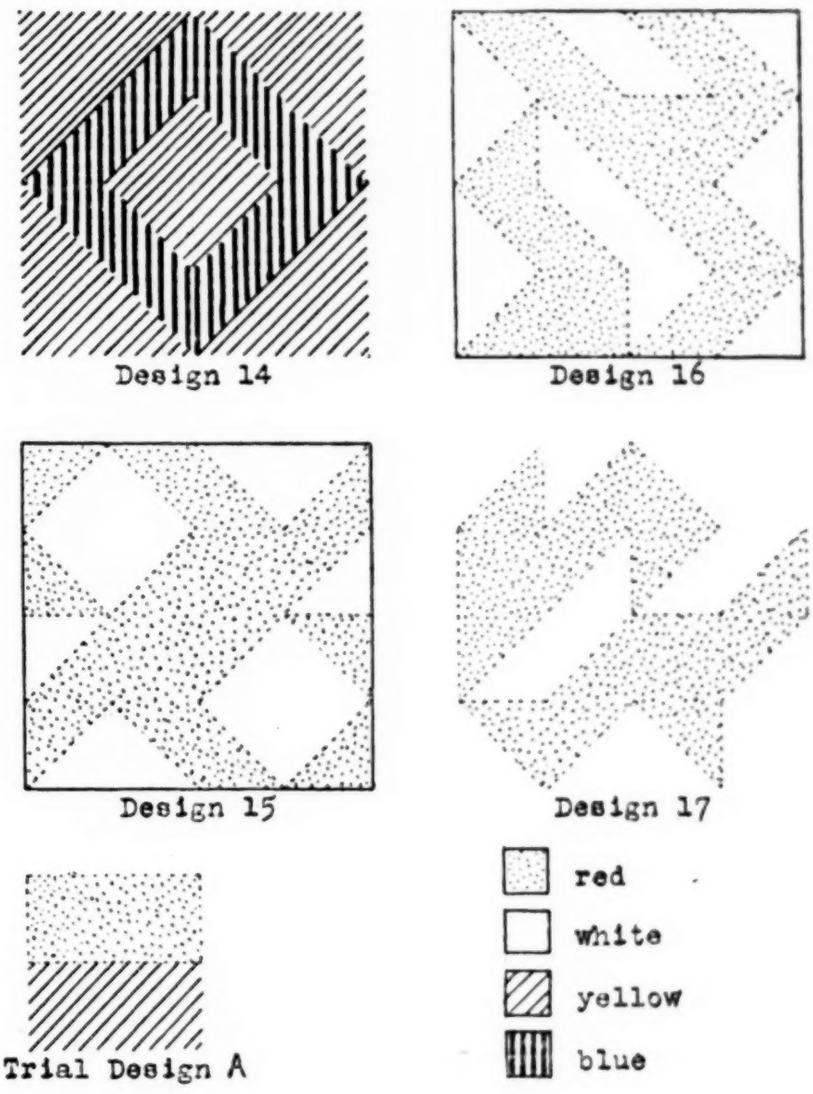


CHART I. Block Designs.

Designs: 1, 2, 3, 4, 5, 6, 7, 8, 9—4 blocks.
Designs: 10, 11—9 blocks.
Designs: 12, 13, 14, 15, 16, 17—16 blocks.
Colors (Windsor & Newton, Ltd.): red = carmine lake; blue = prussian blue; yellow = pale cadmium.

To perform the test, utilizing twenty designs, one averaged about an hour or an hour and a half. In the final revision three designs have been eliminated, leaving seventeen, thus decreasing somewhat the time necessary to apply the tests with no significant decrease of reliability. The criteria for rejection were based on correlations with those arrays of evidence presuming to yield an index of intelligence, such as is

obtained through the use of the Binet Scale, as also upon the basis of the diagnostic value of each design determined by the progress of its curve with increasing chronological age. The results at present indicate that the block designs are as good as any single test in the Binet scale (though better in the sense of *diagnostic value*), as good as the Trabue Language Completion Tests, or any other similar single type test, whether involving the use of language or whether mere performance.

The designs, appropriately colored, are printed on medium-thick, white, semi-gloss cardboard. The dimensions of the card are 3 by 4 inches. The printed designs, placed in the center of the card, are *one fourth* the size of the actual designs when the cubes are used. In other words, the face of a cube represented on the designs is only *one half* of an inch on each of its sides. Thus design no. 1 is one inch square, design no. 10 is one and a half inches square, and design no. 14 is two inches square.

The writer has found it of assistance to place in the lower right-hand corner the time limit for each design. These values follow:

TABLE I

TIME LIMITS FOR EACH DESIGN

Design (Number)	Time Limit (Minutes)	Design (Number)	Time Limit (Minutes)
1	1½	10	3
2	1½	11	3½
3	1½	12	3½
4	2	13	3½
5	2	14	3½
6	2	15	4
7	2	16	4
8	2	17	4
9	2		

The time limit as set for each design is about one minute longer than the time within which a correct response may reasonably be expected.

It may be of interest to remark that if the full limit is allowed on each test the working-time totals only 45 minutes for all of the seventeen designs. With practice an examination should average about thirty to forty minutes. In some

cases it may take only fifteen or twenty minutes, in others perhaps an hour.

(*B*) The Directions for Applying the Tests

Preliminaries

Seat the subject comfortably at a table, noting that his visual angle when working with the tests is not less than 45 degrees. Be sure that no designs are visible in your preliminary instructions, nor more than a single design at any one time. The blocks which are not being utilized should be kept in a box, apart, so that they are either invisible to the subject, or if visible, the blocks should be arranged so that the top sides are all of the same color.

Method: Part 1. For Subjects Who Can Understand Spoken Language

(Section *A*) Take a block. (Instructions to subject are placed in quotation marks. For Design 1 four blocks will have been removed from the box.) "Here are some blocks,—give me the name of the color on this side." Sides with the full color are presented first. Place your finger on the side designated. After the subject has responded, turn to another side. "And what is the color on this side?"—"Now the color here?"—"And what is the color here?"—If the subject has succeeded in naming the colors correctly, proceed with the experiment. (If he has failed, further instructions are given below, in Part 2.) Then the experimenter explains: "Now on this side we have blue and yellow, (point), and on this side red and white (point). And all the blocks are painted in the same way."

(Section *B*) "What you are to do is this: Take these blocks," (Shuffle them so that when finally placed before the subject, no more than one fourth of the blocks have topside colors which are present in the design, the separate blocks being placed apart, flat on table, and not piled one on top of another) "pick out the right colors, put them together, and make them look, on top, just like this." (Point to design 1.) Give no further hints nor suggestions if the directions have

been understood. CAUTION: Be sure that all the blocks are thoroughly shuffled *before* the design is presented. The purpose is to eliminate the possibility of studying the design before being ready to begin work with the cubes.

(Section *C*) If the subject has not understood what is meant, the experimenter may perform trial design (*A*)[1] slowly, using pantomime freely, the subject watching closely, after which the subject is requested to repeat the operation. This may be repeated any number of times until the subject understands. When he does, proceed with the designs in order, beginning at (Section *B*), and continuing with (Section *D*).

(Section *D*) After the first design has been completed or failed, the blocks are again shuffled, observing the cautions in (Section *B*), and the subject is told again to "Take these blocks, pick out the right colors, put them together, and make them look on top just like this." (Point to design number 2.) The instructions remain the same for all the designs. The subject is not told at any time the number of the blocks he is to use.

Record.—Both time and moves are recorded. A move is counted when a block is given its initial position on the table. Each separate and distinct change in the position of a block is counted a move. Sometimes a child will make three or four changes in the position of a cube, the topside remaining the same color (especially true of diagonal sides, *e.g.*, red-white). But each change in position is counted a separate move. If success is not attained within the time limit, no credit is assigned. The time limits are indicated on the design cards.

The whole test is not regarded as complete unless there are, ordinarily, at least five consecutive failures on designs after the last success, and where doubt exists as to inability in the later designs, give as many designs beyond the last success as is deemed wise.

[1] Trial Design (*A*) is represented on the pages with the other designs and is used only when under the provisions of Section *C* further preliminary explanation is necessary. Trial Design *A* is a four-block design, two full red sides above, two full yellow sides below.

Part 2. For Subjects Who Do Not Know the Names of the Colors

Take all the blocks out of the box and place on the table so that the single-colored faces are all on the top side of the cubes. Have an equal number of reds, yellows, blues and whites. Point to a red-topped block and ask the child to point to all the blocks that have the same color on top. Do the same for the other three colors. If the child can distinguish the colors, proceed with the test at (Section *B*).

Part 3. For Subjects Who Cannot Understand Spoken Language

By means of gestures and pantomime go through the procedure in Part 2. If the subject can distinguish the colors, proceed with (Section *C*), and through the various designs. The method of recording remains the same.

(*C*) THE SCORE CARD AND THE METHOD OF SCORING

In the following table are presented the score values of each of the seventeen designs and the number of score points to be deducted if a design is successfully completed with excess time and with excess moves:

TABLE II

SCORE CARD

Design No.	Score Value	Points to be Subtracted		
		Time		Moves
		1 Point	2 Points	1 Point
1	3	21″ and over	——	6 and over
2	5	31″ and over	——	7 " "
3	6	21″ to 35″	36″ and over	8 " "
4	6	31″ to 1′ 0″	1′ 1″ " "	10 " "
5	7	36″ to 1′ 5″	1′ 6″ " "	11 " "
6	7	36″ to 1′ 0″	1′ 1″ " "	12 " "
7	7	41″ to 1′ 10″	1′ 11″ " "	11 " "
8	8	41″ to 55″	56″ " "	10 " "
9	9	56″ to 1′ 10″	1′ 11″ " "	15 " "
10	9	1′ 56″ to 2′ 10″	2′ 11″ " "	22 " "
11	8	1′ 46″ to 2′ 30″	2′ 31″ " "	19 " "
12	9	2′ 26″ to 2′ 40″	2′ 41″ " "	30 " "
13	9	2′ 21″ to 2′ 33″	2′ 34″ " "	31 " "
14	9	2′ 26″ to 2′ 40″	2′ 41″ " "	32 " "
15	9	2′ 41″ to 3′ 0″	3′ 1″ " "	32 " "
16	10	2′ 41″ to 3′ 5″	3′ 6″ " "	31 " "
17	10	2′ 41″ to 2′ 55″	2′ 56″ " "	30 " "

Maximum score—131 points.

To clarify the table, one or two illustrations will be utilized. For example, design number two has a score value of 5. This full amount is attained if a reagent completes the design successfully in less than 31 seconds and with less than 7 moves. If 31 or more seconds are utilized, one point is deducted from the score, and if 7 or more moves are made an additional point is deducted. Take again design number thirteen which has a score value of 9. This full amount is attained if the subject completes the design successfully in less than 2 minutes and 21 seconds, and with less than 31 moves. If completed between 2 minutes 21 seconds and 2 minutes 33 seconds, one point is deducted, if 2 minutes 34 seconds or more are spent on the problem, two points are deducted. And if 31 or more moves are made an additional point is deducted from the score value of the design.

The scoring of a performance is a very simple matter. This will be self-evident from the following examples:

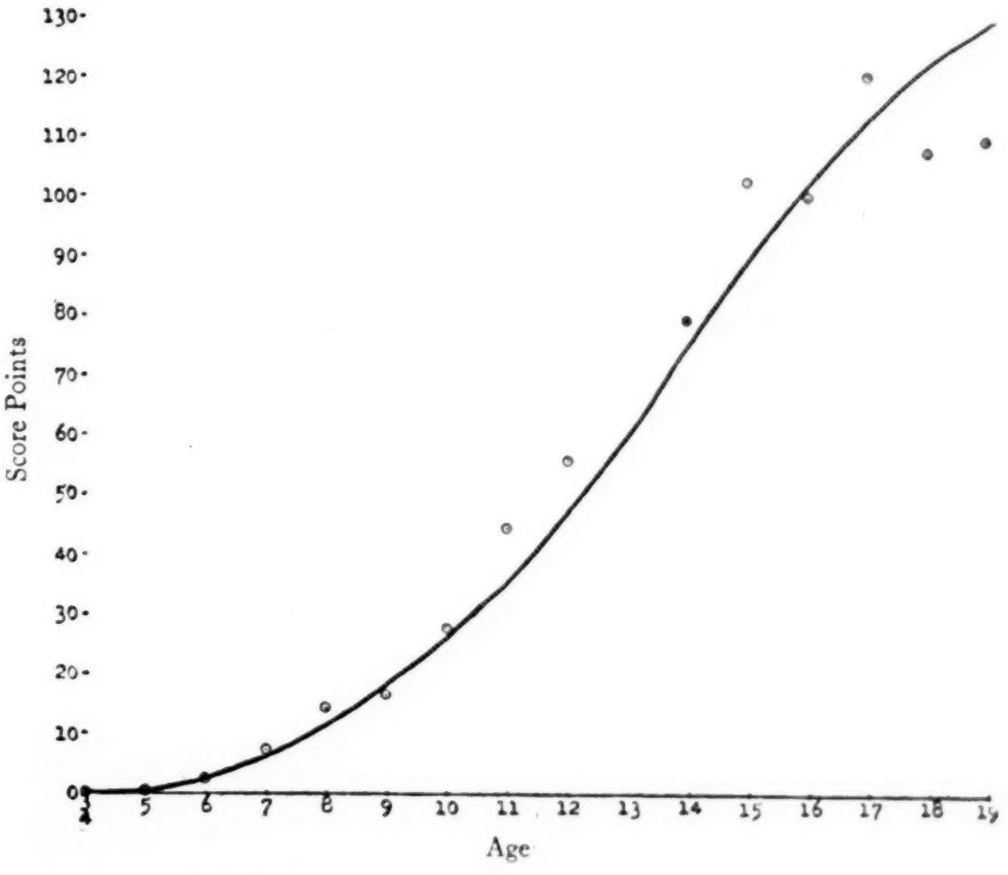
Example one: Design number 7 successfully completed in 1 minute and 23 seconds and at the end of 9 moves. Score 7, for successful completion, less 2 points for excess time. Final score 5. *Example two:* Design number 10 successfully completed in 1 minute 48 seconds, and after 19 moves. Score 9, for successful completion. No deductions for excess time or excess moves. Final score, 9. *Example three:* Design number 16, successfully completed in 3 minutes 27 seconds, and after 48 moves. Score 10, for successful completion. Deduct 2 points for excess time, and one point for excess moves. Final score 7.

It may be worth remarking that successful performance, speed and what may be termed accuracy are all combined in the final score. Successful performance receives greatest weight, speed next and accuracy next. The weight ratio as explained elsewhere in the monograph is roughly 4 : 2 : 1. This ratio has been empirically determined and was not derived by arm-chair philosophizing. The prevalent opinion, which was at one time shared by the writer, that speed and accuracy cannot be combined in one score, does not hold with the Block Design Tests. The writer felt that success, speed

and accuracy each had its own diagnostic importance and in order to make the tests most effective all should and must be taken into account in the final score summation. But more of this in the longer monograph.

(*D*) The Norms

The procedure involved in obtaining norms for the different designs was quite a complicated one, requiring a great deal of careful statistical work. In this effort the writer utilized the currently accepted standardization methods, with but slight modification. An explanation of the general pro-



Graph I. Mental Age Equivalents of Score Points.

cedure utilized together with a description of various methods of checking the results has been left for the later monograph. Suffice it to say that the score points mentioned in Table II. are to be interpreted in the same light as those of Buckingham in his standardization of his spelling tests, of Trabue in his

standardization of his language-completion tests, and of Woody in his standardization of his arithmetic tests. In this section the final results, merely, will be presented.

Graph I. is the curve indicating the scores to be expected at the various ages from 3 years to 19 years. This curve has been smoothed but slightly within the range of ages below ten, though rather considerably from fifteen to nineteen. This was necessarily the result of a deficiency in the number of cases at the higher ages. The median score at each age is represented by a circlet with a dot enclosed.

TABLE III

MENTAL AGE EQUIVALENTS OF SCORE VALUES

Score Points	Mental Age	Score Points	Mental Age	Score Points	Mental Age	Score Points	Mental Age
0	5-3 or below	33	10- 9	66	13- 5	99	15- 9
1	5- 7	34	10-10	67	13- 6	100	15-10
2	6- 0	35	10-11	68	13- 6	101	15-11
3	6- 3	36	11- 0	69	13- 7	102	16- 0
4	6- 6	37	11- 1	70	13- 8	103	16- 1
5	6- 9	38	11- 2	71	13- 9	104	16- 2
6	7- 0	39	11- 3	72	13- 9	105	16- 3
7	7- 3	40	11- 4	73	13-10	106	16- 4
8	7- 6	41	11- 5	74	13-11	107	16- 5
9	7- 8	42	11- 6	75	14- 0	108	16- 7
10	7-10	43	11- 7	76	14- 1	109	16- 8
11	8- 0	44	11- 8	77	14- 1	110	16- 9
12	8- 2	45	11- 9	78	14- 2	111	16-10
13	8- 4	46	11-10	79	14- 3	112	16-11
14	8- 5	47	11-11	80	14- 4	113	17- 1
15	8- 7	48	12- 0	81	14- 5	114	17- 2
16	8- 9	49	12- 1	82	14- 6	115	17- 4
17	8-10	50	12- 2	83	14- 7	116	17- 5
18	9- 0	51	12- 3	84	14- 7	117	17- 6
19	9- 1	52	12- 4	85	14- 8	118	17- 8
20	9- 3	53	12- 5	86	14- 9	119	17- 9
21	9- 4	54	12- 6	87	14-10	120	17-10
22	9- 6	55	12- 7	88	14-11	121	18- 0
23	9- 8	56	12- 8	89	15- 0	122	18- 2
24	9- 9	57	12- 9	90	15- 0	123	18- 3
25	9-11	58	12-10	91	15- 1	124	18- 5
26	10- 1	59	12-10	92	15- 2	125	18- 7
27	10- 2	60	12-11	93	15- 3	126	18- 9
28	10- 3	61	13- 0	94	15- 4	127	18-11
29	10- 4	62	13- 1	95	15- 5	128	19- 1
30	10- 5	63	13- 2	96	15- 6	129	19- 3
31	10- 7	64	13- 3	97	15- 7	130	19- 7
32	10- 8	65	13- 4	98	15- 8	131	19-11

In Table III. are presented the mental age equivalents of each score from 1 (mental age 5 years 7 mos.) to 131 (mental age 19 years 11 mos.)

(*E*) Reliability of the Tests

To measure the reliability of any newly devised test of intelligence is not a simple matter. It devolves upon the standardizer to present evidence that the new intelligence scale measures this inadequately defined entity 'intelligence' with approximately the same degree of accuracy as those standards or measuring 'rods' now commonly accepted and in current use.

In this brief article the writer will limit himself to five criteria:

(1) The mental processes employed;
(2) Increase in score from year to year;
(3) Correspondence of median mental ages;
(4) Correlations between mental ages, intelligence quotients and teachers' estimates of intelligence;
(5) Conformance of intelligence-quotient distribution with normal probability.

(1) *Mental Processes Employed*

In devising and standardizing this test the writer did not approach the problem with any bias of 'faculty psychology.' The idea still seems prevalent, though not as much now as in the immediate past, that in order to possess an adequate measuring instrument for intelligence, the device must contain separate tests for each mental 'function': sensation, perception, association, imagination, memory, judgment, reasoning, etc. On the other hand it has been amply demonstrated that the only intelligence scales worth the name draw service freely from all 'functions.' Binet has pointed out that all 'intelligent' operations involve the functioning of three primary activities: first, *attention* to the problem presented; second, a conscious attempt on the part of the subject to consummate an adequate *adaptation* to the situation; and third, the exercise of *auto-criticism* in order to determine how efficiently the specific 'adaptation' has solved the problem. A cursory examination of the demands made upon the mental operations of the person tested with the block-designs will clearly reveal that *attention*, *adaptation* and *auto-criticism*

are all involved in the successful accomplishment of each task. That point in the graded series of designs at which a child will begin failing to achieve further success, will be a rough measure of the development of his ability to attend, to adapt and to critically survey his general plan of performance and his ultimate accomplishment. In his discussion of the 'patience test' in the 1908 scale, and these words might as well apply to the block-design tests, Binet states:[1] "It is a game, but at the same time a work of the intelligence. When one analyzes the operation it is found to be composed of the following elements: (1) Consciousness of the end to be attained, that is to say, a figure to be produced; this end must be understood, and kept in mind; (2) the trying of various combinations under the influence of this directing idea, which often unconsciously determines the kind of attempt which should be made; (3) judging the combination formed, comparing it with the model, and deciding if it resembles the other" (p. 198). If 'intelligence' involves the following mental operations; analyzing, combining, comparing, deliberating, completing, discriminating, judging, criticising and deciding, then the block-design tests may, with justice, be said to call upon the functioning of intelligence and to that extent they are a measure of that mental capacity.

(2) *Increase in Score from Year to Year*

As regards the second criterion, reference to Graph I. and to the various tables presented in this article will clearly demonstrate that this requisite is satisfied. The following, however, should be mentioned: At each life age a greater scatter or range in ability is noticeable than is the case with the Binet tests. Whether this phenomenon argues for reliability or not is left for discussion in the later monograph.

(3) *Correspondence of Median Mental Ages*

At each life age do the median mental ages obtained by the block-design tests correspond with the median mental ages obtained by the Binet tests? This question is an important one, and the extent of correspondence or deviation

[1] The Development of Intelligence in Children, Publication No. 11, Vineland, 1916.

should measure very largely the reliability of the newly devised tests.

In the following table this comparison is presented:

TABLE IV

CORRESPONDENCE OF MEDIAN MENTAL AGES

Life Age Yrs.	No. of Cases	Median Binet Age Yrs.-Mos.	Median Block-Design Age Yrs.-Mos.	Difference Between Medians (Mos.)	Average of Two Medians Yrs.-Mos.
6.....	16	6- 1	5- 3	10	5- 8
7.....	27	7- 5	6- 8	9	7- ½
8.....	28	8- 0	8- 2	2	8- 1
9.....	27	9- 0	8-11	1	8-11½
10.....	33	9- 8	10- 3	7	9-11½
11.....	30	10- 6	11- 6	12	11- 0
12.....	29	11-10	12- 5	7	12- 1½
13.....	32	11- 5	12- 6	13	11-11½
14.....	28	13- 9	13- 6	3	13- 7½
15.....	19	13- 3	14- 0	9	13- 7½
16.....	16	15- 6	13-10	20	14- 8
				Average—8.5	

Four important items are worthy of note: In the first place, the average deviation of the median Binet ages from the life ages at each year is 6 months; second, the average deviation of the median block-design ages from the life ages at each year is 8.8 months; third, the average deviation between the two intelligence-test medians is 8½ months, and finally, the arithmetic mean of the two medians for each life age results in a more accurate approximation of what may be the 'true' mental age than either median taken alone. The significance of the last fact will have to be left for more complete discussion in the later monograph. At this point, it may be sufficient to remark that the approximation between the Binet and block-design medians is rather close, especially when we consider that the block-design tests are quite free of the 'language factor.'

(4) *Correlations between Mental Ages, Intelligence Quotients and Teachers' Estimates of Intelligence*

In order to understand and to justly evaluate the relations about to be presented, the Binet results will be mentioned to serve as a standard of comparison.

1. The correlation[1] between Binet age and life age is + .80 (P.E. ± .01) (291 public school cases).

2. The correlation between block-design age and life age is + .66 (P.E. ± .02) (291 public school cases).

3. The correlation between Binet age and block-design age is + .82 (P.E. ± .01) (366 cases). The table is herewith presented:

TABLE V.

BINET AGE AND BLOCK DESIGN AGE

Block-Design Age	*Binet Age*																	
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
Below 5	3	7	14	14	3	1		1										43
6				10	10	10	2	3	3	1	1							40
7				2	7	9	7	2	1			1						29
8				4	3	4	10	4	2		2	1						30
9				1	3	14	6	7	1	4	2							38
10					2	6	5	8	5	2	3		2					33
11						6	2	4	9	4		1		1				27
12					1	1	2	6	8	3	4	2	1					28
13							1	3	7	5	3	1	1	1				22
14							2	3	2	7		4		2				20
15								2	3	5	3	4	1					18
16								1	2		1	1	1	2		1	3	12
17									1			2	4	1	3	2	1	14
18											1				2			3
19													1	2			1	4
20														2	2	1		5
Total	3	7	14	31	29	51	37	44	44	31	20	17	11	11	7	4	5	366

(Note: Age 10 means 10–7 to 11–6 etc.)

4. The correlation between Binet age and block-design age is + .81 (P.E. ± .01) (291 public school cases).

5. The correlation between Binet age and block-design age is + .67 (P.E. ± .05) (75 feebleminded cases).

6. The correlation between Binet I. Q. and block-design I. Q. is + .80 (P.E. ± .01) (366 cases). The table is herewith presented:

[1] Pearson's Coefficient (*r*) used throughout.

TABLE VI

BINET I. Q. AND BLOCK-DESIGN I. Q.

Block-Design I. Q. \ *Binet I. Q.*	20	30	40	50	60	70	80	90	100	110	120	130	140	Total
30		13	7	1	1	1								23
40	1	3	4	6	2	3		1						20
50		1	1	5	7	1	1	2						18
60			2	2	4	8	10	1	2					29
70				1	1	4	6	4	3	2				21
80					1	3	11	12	10	5	2			44
90				1	1	3	9	10	20	2	2			48
100						1	6	20	20	7	1	1		56
110						2	7	10	13	7	2	1		42
120							1	5	8	8	5	1	1	29
130								1	3	4	5	2		15
140								2	3	3	5	1	1	15
150									1	1				2
160										1	1		1	3
170											1			1
Total	1	17	14	16	17	26	51	68	83	40	24	6	3	366

(Note: I. Q. of 50 means 46–55, etc.)

7. The correlation between Binet I. Q. and block-design I. Q. is + .57 (P.E. ± .03) (291 school children).

8. The correlation between Binet I. Q. and block-design I. Q. is + .67 (P.E. ± .05) (75 feebleminded cases).

9. The correlation between teachers' estimates of intelligence and Binet I. Q. is + .47 (P.E. ± .03) (291 school children).

It may be worth remarking that although the correlation between block-design age and Binet age is + .82, teachers' estimates of intelligence correlate only one half as much with the Block-design I. Q.'s as with the Binet I. Q.'s. The reader may recall that one of the original objections to the Binet scale was that it measured school training. Only to a limited extent has this been denied, the explanation having been made that the tests measure intelligence through the medium of knowledge only partly influenced by school training. It has been admitted, true, that practically all children are exposed to these educational influences, but the ultimate difference in achievement is explainable on the basis of differences in endowment. However this may be, the results of the block-design test would perhaps tend to show that there is more to this charge than we have been inclined to admit.

It will, no doubt, be acceded without much question that the block-design tests are less affected by school training than the Binet.

At any rate the total correlational evidence seems to indicate that the block-design tests possess a high degree of reliability.

(5) *Conformity of Intelligence Quotient Distribution with Normal Probability*

A very necessary index in weighing the reliability of any standardized test is to determine the extent to which an actually found distribution conforms to its theoretical distribution.

In the following table are presented the I. Q.-range-distributions for the Binet and the block-design tests. The respective percentage values are compared with what one should theoretically expect.

TABLE VII

INTELLIGENCE QUOTIENT RANGES

	26 to 35	36 to 45	46 to 55	56 to 65	66 to 75	76 to 85	86 to 95	96 to 105	106 to 115	116 to 125	126 to 135	136 to 145	146 to 155	156 to 165	166 to 175
Stanford Binet Obtained	....	.....	....	1.7	5.5	16.5	22.7	28.2	13.8	8.3	2.1	1.0	....	....	....
Theoretical Expectation	....	.....	....	.16	1.6	8.5	23.42	32.64	23.42	8.5	1.6	.16	....	....	....
Block-Design Obtained	.034	.034	1.4	6.5	6.2	14.4	15.1	18.9	14.4	10.0	5.2	5.2	.07	1.0	.034
Theoretical Expectation (Median at 99)...	.49	1.28	2.78	5.21	8.67	12.05	14.66	15.30	13.84	10.69	7.25	4.11	2.08	.88	.33

The average deviation from theoretical expectation for the Binet I. Q. ranges is 3.3 per cent. per I. Q. group. The average deviation for the block-design tests is only 1.4 per cent. per I. Q. group.

In conclusion, one may state that the evidence presented seems to indicate not only that the tests measure intelligence,

but that this is accomplished with a fair degree of accuracy. On the other hand, one should bear in mind Stern's caution:[1] "Psychological tests *must not be overestimated*, as if they were complete and automatically operative measures of mind. At most, they are the psychographic minimum that gives us a first orientation concerning individuals about whom nothing else is known, and they are of service to complement and to render comparable and objectively gradable other observations—psychological, pedagogical, medical—not to replace these."

(*F*) Serviceability

In his 'Stanford Revision of the Binet-Simon Scale' (Warwick and York, 1917) Terman states (p. 150) that "to be widely serviceable a test should demand only the simplest material or apparatus, should require at most but a few minutes of time, and should lend itself well to uniformity of procedure in application and scoring." The writer has attempted to satisfy these demands in standardizing the block-design tests. Those who utilize the tests will find after a little practice that there can be but little variation in the findings of two examiners, and that the only chance for difference is in the recording of the number of moves made.

The special value of the block-design tests lies in the fact that valid results may be obtained independently of the 'language factor.' Neither deafness nor lack of language understanding should be disqualifications in the proper performance of the test. The block-designs may therefore be utilized in the study of racial differences, in determining the mental capacities of the deaf and of those suffering from various other language handicaps.

As regards the borderzone problem, although further investigation of this matter by the writer is now under way, it seems that this test will aid in a better differentiation of the group of cases falling in this category. The writer maintains that feeblemindedness is *not* an arbitrary statistical designation, but is rather a clearly demarked physiological entity

[1] 'The Psychological Methods of Testing Intelligence,' Warwick and York, 1914, p. 12.

quite distinct from normality, statistical-psychologists notwithstanding. Years of experience with this type of defect has fixed the notion in the writer's mind that feeblemindedness is indicative not only of mental mal-functioning, but also of physiological mal-functioning, especially of endocrine character. The results of further research, however, can be the only tests of the truth of one's statements at this time.

The writer regrets the omission of much pertinent material in this brief presentation, but the later monograph will deal with many topics here barely touched upon, if at all.

Addendum

Regarding the Average Mental Age of Adults

Of importance in interpreting the results of this newly devised mental test is the recently raised question regarding the average mental age of adults. In the promised monograph a few pages will be devoted to a psychological and statistical discussion of this important matter. At this point the writer merely wishes to state that the data so far presented does not warrant accepting the suggestion that "The previous notion that the average intelligence of adults is 16 years should be given up."[1] There is a fundamental fallacy underlying the suggested 13 to 14 year criterion, a complete discussion of which must be left for a later time.

[1] E. A. Doll: New Jersey State Prison, Psychologist's Report (1918–1919), p. 72.

RECENT APPARATUS FROM THE PSYCHOLOGICAL LABORATORY OF McLEAN HOSPITAL

F. L. WELLS, PH.D., AND C. M. KELLEY, M.D.

I. Galvanometer control board 377
II. Fall apparatus 378
III. Choice-reaction switchboard 381
IV. Fall screen 383
V. Multiple-pen chronograph 385
VI. Device for water-recording of motor tests 388

1. A Galvanometer Control Board

1. The functions of the control board are threefold: to provide through a shunt resistance and a galvanometer (1) a circuit from psychogalvanic electrodes, (2) a current to another circuit such as for psychomotor reaction times, or (3) a known resistance to the current.

2. The units of the board are, with one exception, contained in a small mahogany-stained cabinet, 16″ x 16¼″ x 9¼″, the cover of which is formed by the forward half of the top, hinged to the other half, and the front side attached to the top by right-angle braces. A primary cell rests in a compartment at one rear corner and fixed to the floor are,—an Ayrton shunt; a known resistance, 5,000 ohms; a single-pole, double-throw switch; a single-throw switch; and six Fahnestock clips grouped in an accessible position. A two-pole, double-throw switch has been placed outside the cabinet, but in future constructions should be within it.

3. For identifying the points of contact in the wiring, let the clips be numbered from 1 to 6 in order of position and the binding posts of the shunt from 7 to 10. A wire from the positive pole of the cell leads to clip 6, which is connected with the forward post of the double-throw switch, and also with one side of the 5,000-ohm resistance. A wire from the other side of the resistance is connected to one post of the single-throw switch, the other post being connected with

clip 3, which in turn is connected with post 7 on the shunt. The negative wire from the cell is attached to clip 5, from which a wire leads to post 8 on the shunt. This post is also connected with the rear post of the double-throw switch, the central post being connected with clip 4. The two central posts of the two-pole switch outside the cabinet are wired to clips 3 and 4, respectively. Wires from the rear posts of this switch lead to the psychogalvanic electrodes; and those from the forward posts, to the reaction-time circuit. The galvanometer wires lead from clips 1 and 2, connected with the shunt posts 9 and 10.

4. The circuit from the psychogalvanic electrodes, excluding the cell, is made by closing both double-throw switches rearward, and the reaction-time circuit by closing both switches forward. Closing the single-throw switch throws the current through the 5,000-ohm resistance, independently of the other two circuits. The shunt and galvanometer, however, remain in each of the three circuits. In psychogalvanic work it is ordinarily desired to use body resistance to the cell current. In this case the double-pole double-throw switch is rearward, the single pole double-throw switch forward.

II. A Fall Apparatus for a Galvanometer Chronoscope

5. A homely but satisfactorily efficient model of a fall apparatus for standardizing galvanometric readings has been constructed in the laboratory of readily obtained and inexpensive materials. The suggestion was derived from 'A New Chronoscope and Fall Apparatus,' by Paul E. Klopsteg (*Journal of Experimental Psychology*, 1917, 2, 253–268). The approximate dimensions of the model are 5′ (height) x 1′ x 8″, and its range, up to 5/10 of one second.

6. The base is a block of oak, 12″ x 8″ x 2½″, provided with three levelling bolts. The upright consists of an iron rod 60″ in length, threaded at its lower end; and two boards, each 58″ x 2½″ x 7/8″, grooved to enclose the iron rod snugly when screwed together. The threaded end of the rod, protruding from the assembled upright, is driven into a

hole through the base and bolted, the upright being further secured in the perpendicular by four right-angle braces. Thus, rigidity is obtained and warping prevented. The upright is graduated by machine-ruled millimeter scales, inserted into the front surface and extending from 0, 3″ from the base, to the top. The respective distances of fall for tenths of a second up to .5″ are marked as nearly as possible to the quantities derived by the formula, $s = \frac{1}{2}gt^2$ (g = 980.6 cm.)[1] as follows:

at	122.575	cm. for	.5″
"	78.448	" "	.4″
"	44.127	" "	.3″
"	19.612	" "	.2″
"	4.903	" "	.1″

7. A freely movable wooden frame, 3″ in height, surrounds the upright and is clamped at the desired level by a thumb-screw through the back, which forces a metal plate against the standard. Four Fahnestock clips are secured to the rear surface of the frame. An electromagnet is mounted on the front surface and is connected by wires to two of the clips. Immediately below the core, but not in contact, are two rigid metal strips with slightly curved ends for contact with a 5/8″ soft-iron ball, a wire from each strip passing to a clip at the rear. The magnet and contact points are so adjusted that the lowest portion of the ball is held tangent to a horizontal line from the lower border of the frame, which accordingly serves as an indicator on the scale.

8. A cradle, supporting a horizontal tray 2″ square with vertical borders, is mounted at the foot of the upright. A slight vertical motion, with the tray constantly horizontal, is attained by four uniform, mutually parallel links pivoted, two on either side, on the wooden block supporting the tray

[1] The acceleration value g (980.6 cm.) is for the latitude of Washington, D. C. Variation due to differences in latitude and elevation is shown by the formula $g = 980.6056 - 2.5028 \cos 2\lambda - .000003h$ (Barker, G. F., Physics, 4th. ed., 1892, 105). At the laboratory, latitude (λ) = 42°21′28″, and elevation above mean low tide (h) = 7620.0 cm. g is found to be 980.353282 cm. This introduces an error of .02516 per cent., which is neglected. For the distance and mass involved, error due to air resistance is also negligible.

and on the upright. The tray is held at rest at 0 on the scale by means of a U-shaped lever, pivoted on each side of the upright, the extremities extending forward under the upper links of the cradle, and the bowl of the U in the rear being drawn down by a spring to rest upon a vertical adjustment screw.

9. Any downward movement of the cradle trips a horizontal trigger by means of an adjustable metal plate on the front. The trigger is a rigid rod supported by three bearings on the end of the wooden base of a telegraph key which is secured to the base of the standard. One end of the rod is bent at a right angle to form a short arm extending under the trip plate, and a slightly longer vertical arm with its extremity hooked and beveled is at the other end. When the key lever, without a knob and with the tip beveled to reduce friction, is depressed it is automatically held closed by the trigger actuated by a low tension spring. The trip plate is adjusted to rest on the short arm when the cradle is at 0.

10. A knife-blade switch, a relay, and two Fahnestock clips (by which the apparatus is brought into the galvanometer circuit) are also attached to the base.

11. The wiring provides for three circuits: (1) the electromagnet; (2) the "ball-contact," through the primary of the relay; and (3) part of the galvanometer circuit. In the magnet circuit, the positive wire from a storage battery is connected to the switch and thence to one magnet clip on the magnet carriage, and the negative wire leads from the other magnet clip to the battery. One primary post of the relay is connected with the positive wire from the battery, the other primary post with one "ball-contact" clip, and the remaining clip with the negative magnet clip. In the galvanometer circuit, one clip on the base is wired to one secondary post of the relay, the other post to one side of the telegraph key, and the other side of the key to the second clip. The contact points of the relay are set to close the galvanometer circuit with the armature released.

12. The procedure of operation consists of the following steps: (1) clamping the magnet carriage at the desired level;

(2) with the trigger key open (and the reaction key on the external galvanometer circuit closed), closing the magnet switch; (3) applying the ball to the contact strips (thereby completing the primary relay circuit and breaking the galvanometer circuit at this point); (4) closing the trigger key (making the galvanometer circuit at that point); and (5) opening the magnet switch, which allows the ball to drop and thus closes the galvanometer circuit by releasing the relay armature. Impact of the ball on the tray releases the trigger key, so that the current flows through the galvanometer only during the time of fall.

13. The galvanometer deflection is read by the excursion of the reflected beam of light on a millimeter scale curved with a radius of 2 m., the focal distance of the galvanometer mirror. The suitable range of deflection is obtained by adjustable resistance in the galvanometer circuit. This circuit is made available for experiments by merely opening the reaction key and closing the trigger key.

14. In his original paper (cf. par. 5), Klopsteg recommends not the use of a relay, but of a high resistance in the galvanometer circuit. The present apparatus is equally adaptable to such a technique. The relay was used here as better adapted to other portions of the set-up and has no practical effect on the reliability of the measurements involved.

III. An Attachment for Choice Reaction Experiments

15. It was desired to provide a signal indicating the correctness of a response or any falsity, either by the wrong key alone or by both reaction keys, in case of either correct or incorrect propositions; and further to provide for breaking the (galvanometer) chronoscope circuit on any response. To meet these requirements, a simply constructed switch-board was devised, operated by a single lateral movement of the hand.

16. The board contains a three-pole circular switch, a relay, and a 6-volt signal lamp, two contact screws being near the left border and three near the right. Wires from the storage battery, on one of which an 'operating key' is

placed, lead to the two left-hand screws; and the three right-hand screws are connected respectively with the outer post of the right-hand reaction key, the inner posts of both keys by a common wire, and the outer post of the left-hand key. The subject is instructed to close both keys on the verbal signal 'set' and to release the right-hand key if the proposition exposed is correct or the left-hand key if incorrect, in either case holding down the other key.

17. The fixed element of the switch is composed of six narrow contact strips attached radially at 60°—intervals on a non-conducting board. They are bent and their edges curved upward to permit sliding plates to pass under them to make contact with slight tension. The moving member is a non-conducting disk, 2″ in diameter, to the lower side of which three independent, sector-shaped plates are firmly attached at 120° intervals between centers. The plates project beyond the disk and under the ends of the contact strips, each connecting two strips. A handle, attached to the upper side of the disk, extends horizontally to the front and its range is limited to 60° by stops. For correct propositions, the disk is rotated to the left and for incorrect, to the right.

18. The board is wired as follows: Numbering the contact points in order of position, point 1 is connected with one wire from the battery; point 2 with one primary post of the relay (the secondary of which is in the galvanometer circuit and is broken while the armature is released); point 3 to the second wire from the battery; point 4 with the 'proper' wire from the right-hand reaction key; point 5 with the second primary post of the relay; and point 6 with the 'proper' wire from the left-hand reaction key. The 'common' wire terminal at the board is connected with one side of the signal lamp, the other side of which is wired to point 2 (connected with the relay primary post).

19. The handle is so placed on the disk that in the 'correct' position, one sliding plate connects points 2 and 3; the second plate, points 4 and 5; and the third plate, points 6 and 1. In position for 'incorrect' propositions, the first plate connects points 1 and 2; the second, points 3 and 4; and the third, points 5 and 6.

20. Tracing the circuits established, it is seen that with the reaction keys 'set,' closing the 'operating' key closes the relay and lights the signal lamp; and breaking contact at any of the keys opens the relay circuit. With the switch in the 'correct' position, releasing the right-hand key (the right response to a correct stimulus) permits the lamp to continue lighted—to brighten, in fact, as the relay magnets are cut out. A false reaction, whether by releasing the left-hand or both keys, extinguishes the lamp by breaking the entire circuit. Switching the handle to the right for incorrect propositions reverses the action of the subject's keys so that releasing the left-hand key (the right response) causes the lamp to brighten, but a false reaction or a reaction with both hands extinguishes the light.

IV. A Fall Screen for the Exposure of Choice-reaction Stimuli

21. This piece of apparatus[1] provides for independent, successive exposures of 25 stimuli and, by connections with a chronoscope and attachments or with a multiple-pen chronograph, for recording the time and accuracy of each response.

22. The standard is of wood, its greatest dimensions being approximately 8′ x 22″ x 18″. The pedestal,[2] comprising 16″ of the total height, consists of a top, 12″ x 17″ x 3/8″, supported by four slightly converging legs of 1¾″ stock and braced by cross pieces. The front of the upright is formed by the floor of a vertical trough 64″ long, 10″ wide, 3½″ deep, and is extended upward by a board 14″ long. A horizontal aperture, 2¾″ x ⅞″, with beveled edges and provided with a cover, is cut 52″ from the floor. The sides of the trough are grooved along their inner surfaces to guide a sliding frame. For 42″ of its length, the trough forms the front of a box of the same width and 9½″ deep, containing a shelf 24″ from the base and 1½″ thick. The rear of the upper compart-

[1] Cf. Wells, F. L. and Sturges, H. A., 'Pathology of Choice Reactions,' *Amer. Journ. Insan.*, 1918, 75, 82–85.

[2] In a later model supplied to the Carnegie Institute of Technology, the upright was, for the greater comfort of the subject, hung at the rear of a table, affording room for the knees.

ment is closed by a cover opening upward, and the lower by a door opening laterally. On the shelf is pivoted a metal disk, 6⅛″ in diameter, from the border of which a radial section 1″ at the circumference and ¾″ deep is cut. A binding post is secured to the disk and a handle is placed diametrically opposite the open section. Thus a hand-operated escapement for the sliding frame is afforded.

23. The sliding frame is a rectangular hollow frame of wood, 30″ x 5–⅝″ x 3″, with a guide strip on either side fitting into the corresponding groove of the trough, and on the rear is firmly attached a row of 27 hard wooden blocks, ¾″ wide, ⅞″ high, and ⅝″ thick, arranged alternately on either side of the middle line (as the black squares in two rows of a checker-board). The lower surface of each block is covered with felt and the blocks are placed with these surfaces at 1-inch intervals. Thus alternate movements of the escapement disk permit 1-inch drops of the frame, moved by gravity. Within the shaft of the frame is a 1½″ brass tube, 43″ long, attached at its base to the pedestal and open at its upper end. A plunger, operating in oil, is connected by an iron rod to a metal arch across the upper end of the sliding frame, whereby the impact of the blocks on the disk is diminished.

24. The stimulus material is printed at 1-inch intervals on interchangeable strips of heavy paper, one of which is slipped down the front of the frame and retained so that one stimulus is exposed at the aperture when the frame is at rest at any station but the first, when blank paper is exposed.

25. A continuous, non-insulated copper wire, arising from a contact clip at the top of the frame, passing over, outside and under each alternate block along shallow grooves, forms contact with the escapement disk except at the instant of change, the break in the circuit being recorded on the chronograph. A flexible wire, on which a weight is suspended by a pulley, is connected between the clip on the frame and another on the shelf. From the latter, a wire leads to one post of a two-pole, single-throw switch on the floor of the box and through the switch, by a 6-conductor Ulesote cable, to the

chronograph. Another wire, attached to the binding post on the disk, passes to the other side of the switch and through the cable to the chronograph circuit.

26. Lateral shelves, 6″ x 10½″, are braced on the sides of the upright 10½″ from the pedestal, a telegraph key being secured to each and concealed, except the operating knob, by a hinged cover, the front of which is open and the border of the top cut away to render the knob accessible without exposing the rest of the key or wiring. The inner posts of both keys are connected by wires to one clip on the pedestal and through the cable to the chronograph circuit. Each outer post, by its individual wire to a corresponding clip and through the cable, is connected with the chronograph electromagnet controlling the pen assigned to the key, so that contact at either key effects its characteristic record.

V. A Multiple-pen Chronograph

27. This is a motor-driven chronograph[1] providing seven simultaneous, independent records, each obtained by a magnetically operated pen striking on moving carbon and paper ribbons against a platen roller. By means of the adjustable speed motor employed and the gearing of the mechanism, any constant rate of motion of the paper between 6″ and 50″ per minute may be obtained. The apparatus is contained in a cabinet, 27″ x 14″ x 13″, the cover of which comprises 10″ of the height and is hinged on the rear wall. Inserted into the top is a glass window, 22½″ x 9¾″.

28. The mechanism is supported by two parallel brass plates, 8″ x 6″, separated by distance rods 2″ long and perforated identically for screws, to provide bearings for the several axles, and by two large openings, to afford accessibility. A rectangular section is cut from one upper corner of each plate to form the seat for a transverse, semicircular frame supporting the magnets, recording points, certain of the rollers, and by a strip of non-conducting composition along the straight border of the upper plate, four binding posts at each end. The frames are mutually attached by small right

[1] This apparatus was built to order by E. S. Lincoln, Inc., of Waltham, Mass.

angle braces, and the two uprights are similarly attached to a base plate, 7⅛″ x 5″, screwed to the floor of the cabinet. The stock is ⅛″ brass plate.

29. A worm-gear box is mounted on the base at the proper level to align the shaft with that of the motor and to enmesh the pinion gear with the main gear wheel in the frame. The ratio of revolutions of the drive shaft and the pinion gear is 40:1. The main gear wheel is 1⅞″ in diameter and is fixed to the axle of the paper take-up drum by a set screw through a collar. It bears 88 spur teeth and the ratio between the pinion, 7/16″ in diameter and having 20 teeth, and the wheel is accordingly 4.4:1. A second pinion gear (11/32″ in diameter; 15 teeth) is fixed on the same axle and is in mesh with an intermediate wheel, 1 11/16″ in diameter with 80 teeth and pivoted on a bar across the upper opening of the supporting plate, the ratio being 5.34:1. On the axle of the carbon ribbon take-up drum, above and to the left of the intermediate wheel, and in mesh with it, is mounted a wheel 3″ in diameter with 144 teeth, revolving on a collar bound to the axle by a set screw and bearing a 12-tooth ratchet. A pawl mounted on the wheel, and provided with a spring, engages in the ratchet to drive the drum but permits a reverse rotation of the drum in rewinding the ribbon. The ratio between the intermediate wheel and the latter is 1.75:1 and accordingly the rate of rotation of the ribbon drum is approximately only 1/9 that of the paper drum (the ratio between the motion of the carbon and paper ribbons being further increased by the difference in circumferences of the drums).

30. The carbon ribbon, 1½″ wide and 8 yards long, arises from a flanged drum of sufficient capacity, fixed to an axle mounted symmetrically with the corresponding take-up drum and extending through the back plate to present a squared end for rewinding with a clock key. A thin strip of brass, attached to one plate, is bent to afford a slight retaining friction on the drum. The carbon ribbon passes over a roller mounted on the transverse frame, between the recording points (below) and the platen (above), and over a second roller to the take-up drum. The dimensions of the rollers

are the same throughout, 1½″ wide between flanges and ¾″ in diameter, and are brass, except the platen which is of composition.

31. The paper reservoir is a drum 2″ in diameter and 1½″ wide (its flanges having a diameter of 3½″), revolving about a detachable axle held in position by two horizontal, slotted arms extending 2½″ from the supporting plates and provided with straight springs, each perforated near the outer end to receive the axle. A heavy brass plate, pivoted at the top of the frame, rests on the roll to serve as a brake. The paper passes upward over a roller parallel to those described but outside and slightly above, between the carbon ribbon and the platen, over a roller similarly placed at the opposite end of the frame, over a third roller half way down the side of the frame, over its take-up drum, between two guide plates, and leaves the cabinet through an opening in its end. The paper take-up drum, operated by the main geared wheel, is 8 9/32″ in circumference and 1½″ wide, being flanged to guide the ribbon. Eight short, sharp pins are driven radially into the surface of the drum at equal distances, ⅛″ from each flange; and an upper and a lower carriage, each provided with two soft felt rollers and held against the drum by a common spring, press the paper onto the pins.

32. The recording device is contained in the transverse frame. At the center of the straight side is placed a row of seven vertical recording rods, tapered and rounded at their upper ends, and each provided with a collar near the lower end. Each rod is freely movable through holes in the horizontal plates but returns by gravity to rest with its collar on the lower plate. Seven electromagnets are mounted between the plates in an arc; and on the border of the lower plate are pivoted the corresponding, converging armatures extending under their respective magnets and their narrowed ends resting under the corresponding rods on a supporting bar.

33. Each group of binding posts is connected by a cable with binding posts passing through the front wall of the cabinet. From the first post on the frame, a common wire

leads to five magnets, the wires from which lead to their respective posts. The sixth and seventh magnets are in separate circuits, their return wires leading to two additional posts in the cabinet wall. The 6th magnet records the incidence of stimuli. The 7th magnet, in circuit with an electric clock, records the time intervals. The other five magnets are placed on reaction circuits.

34. The motor was constructed by the Holtzer-Cabot Electric Company. It is a direct current, adjustable speed, shunt-wound motor; type HD; size 14; H.P. 1/10; 110-volt; 1.3 amperes; with a speed rating of from 300 to 2,400 r.p.m. The control is placed for the convenience of the experimenter; preferably the recording apparatus itself should be placed in a room other than that of the experiment because of the sounds produced by the action of the recording points.

VI. A Device for Water Recording of Motor Tests

35. The material for the contrivance consists of two wide-necked bottles of 150 c.c. and 500 c.c. capacity, with tightly fitting rubbed stoppers; glass tubing; rubber tubing; a glass funnel; an atomizer bulb; two valves from a second bulb, and a spring clamp.

36. The larger bottle (*A*) is upright and its stopper is perforated by three holes, through one of which the stem of the funnel is inserted. Over the end of the stem is slipped the end of a short piece of rubber tubing containing a valve closing upward. Through the second hole, a glass tube extending to the bottom of the bottle is inserted and its outer end connected with the second bottle by rubber tubing which contains a valve closing downward. A short glass tube is placed in the third hole and its free end connected by rubber tubing with the bulb. A short pin is pressed into the rubber plug of the bulb valve to facilitate releasing the air pressure.

37. The smaller bottle (*B*) is held inverted at a convenient height by a standard, its stopper having four holes for as many glass tubes. The shortest tube (inlet) extends merely through the stopper and to it the tubing from bottle (*A*) is connected. The second glass tube (outlet) extends slightly

higher into the bottle and at its lower end, another tube of small calibre is connected by a short rubber tube provided with a clamp. The last-named glass tube is held vertical by an arm from the standard. The third hole of the stopper contains a glass tube (overflow) of sufficient length to permit adjustment and at its lower end, rubber tubing leads to the funnel into which it empties. The tube through the fourth hole extends the length of the bottle permitting the free passage of air.

38. Bottle (*A*) is nearly filled with water and compressing the bulb, by increasing the air pressure within the bottle, forces water into bottle (*B*), its return being prevented by the valve. The volume desired for the experiment is determined by the level of the overflow tube above the outlet tube, the excess returning to the funnel of bottle (*A*), into which it passes on the release of outward pressure against the valve. The adjustment for volume having been made, the procedure of the experiment consists merely of compressing the bulb, opening the clamp to expel air from the outlet tube, and after reclamping the tube, compressing the bulb until water flows into the funnel. After a moment's pause until the excess is wholly drained off, on the release of the clamp the predetermined volume of water falls from the outlet.

39. A simple application of the device is in tests of motor steadiness or tremor where it is seen how much of the predetermined volume of water, issuing from the small outlet tube, the subject can catch in a vessel with its inlet slightly larger than the outlet tube.

SEX DIFFERENCES IN THE EFFECT OF DISCUSSION

BY HAROLD E. BURTT
Ohio State University

Introduction

The following results were obtained incidentally in connection with a study[1] of the inspiration-expiration breathing ratio as diagnostic of truth and falsehood. In many of the series a 'jury' was present primarily to enhance the emotional state of the subject but a number of interesting problems presented themselves in connection with the judgments made by the jury. The principal one was suggested by Münsterberg's brief experiment[2] in which a group of men discussing an objective situation (number of geometrical figures on cards) profited much more by the discussion than did a corresponding group of women. The writer was a member of the Harvard seminar that furnished the male results and was also familiar with the type of students in the Radcliffe seminar and has felt that the groups were not of the same caliber. The women were for the most part an unselected group of college undergraduates and the men a selected group of graduate students. The present situation afforded an opportunity to check this experiment on a much larger number of people with both sexes of about the same academic status. A minor problem that suggested itself was the general ability to judge a person's veracity by his casually observable behavior (intuition) and especially sex differences in this respect. A further problem related to the ability to judge the veracity of a person of one's own sex as compared with a person of the opposite sex.

Method

The subject was given two papers face down, marked on the backs respectively T and L. If he (or she) wished to

[1] To be published shortly.

[2] Münsterberg, H., 'Psychology and Social Sanity,' 1914, pp. 181–202.

tell a lie on that trial he selected the paper marked L and did not touch the other. The paper recounted in considerable detail an imaginary crime in which the subject was implicated. After four minutes study he was questioned on at least ten crucial points and lied consistently to clear himself. If he selected the paper marked T he was supposed to be charged with the same crime but to be innocent and the paper afforded him the details of a complete alibi. He merely recounted these in response to the same ten or more questions. The L was prepared by the experimenter and the T by an assistant so that during an experiment no one but the subject knew whether he was following the alibi provided or making it up himself. The subject wore a pneumograph and sphygmomanometer. The former recorded the inspiration-expiration ratio before and after the replies to questions and the systolic blood pressure was taken at various points during the examination.

In all the series involved in the present study a number of persons sat in front of the subject. There were usually 'jurors' of both sexes present. The nature of the experiment was explained to them and they were then instructed to "watch him closely and try to decide whether he is using the T or the L paper, *i.e.*, whether he is using the alibi prepared by the assistant or making up his own story."

Immediately after the examination a ballot was taken and the result announced. Then the jury discussed the matter openly for about five minutes. The experimenter did not enter into the discussion except when someone was very palpably judging by illegitimate criteria. The discussion usually centered around the person's behavior, nervousness or calm, position of the eyes, facial expression, rapidity of response and sometimes such things as whether the subject had had sufficient social experience or was clever enough to devise such an alibi. After the discussion a second ballot was taken.

No effort was made to duplicate actual legal conditions. The jury was not weighing evidence to any considerable extent but rather deciding upon the veracity of the witness.

Furthermore the subjects, while not instructed to do so, made every effort to deceive the jury and the experimenter. This 'set' was very pronounced in the experienced subjects and involved simulating guilt while actually telling the truth and vice versa, whereas in real life the witness is practically always simulating a truthful attitude.

Two series of this sort, *i.e.*, two "crimes," were performed during an hour's sitting. The subject was privately instructed to select one T and one L in whichever order he pleased. This was to insure an equal number of series of each sort. The jury did not know this, however, but were told the subject would do whichever he wished. This enabled them to judge each series on its own merits.

The experiments were performed in the psychology laboratory of Ohio State University in the spring of 1920. There were in all 22 series performed, 11 T and 11 L, and the total number of jurors was 244. Of these 88 were women and 156 men. They were present in groups of from 4 to 26. Three subjects—2 women and 1 man—participated in the experiment weekly and had had some experience prior to the series with the jury. Four other subjects—1 woman and 3 men—performed two or more series without previous experience in the experiment.

The Effect of Discussion

There are two ways of evaluating the effect of discussion on accuracy of judgment—by comparing the per cent. of the total number of judgments that are correct before and after discussion or by noting the per cent. of the cases in which the judgment is changed in the right or wrong direction as a result of the discussion. These results are given in Table I.

Table I

Per Cent. Correct Judgments

	Female Judges	Male Judges
Before discussion	48	48
After discussion	52	47

Per Cent. Changes in Judgment

	Female Judges	Male Judges
In right direction	9	5
In wrong direction	5	6
Total	14	11

The table shows that 48 per cent. of the women's initial judgments are correct and 52 per cent. of their final judgments. The corresponding figures for the men's judgments are 48 and 47 per cent. This apparently indicates mere guesswork in both judgments and an absence of any effect of the discussion. The lower part of the table shows however that discussion does play a part but works both ways. The women change an incorrect judgment before discussion into a correct judgment after discussion in 9 per cent. of the cases and change a correct judgment to an incorrect in 5 per cent. The corresponding figures for the men are 5 and 6 per cent. Thus while judgments are *altered* by the discussion they are not on the average *improved*. The sex differences in this respect are negligible. In most cases the jury comprised persons of both sexes. In one jury composed of 6 women the initial and final judgments are the same, and in another composed of 26 men there are two changes in opposite directions thus neutralizing each other in the totals.

Münsterberg (using about 15 judges of each sex) found both initial and final judgments for the women 45 per cent. correct, but for men the initial 51 and the final 78 per cent. correct. The present study with 88 female and 156 male judges fails to find such a result. The experiments seem sufficiently similar to warrant a comparison. In both instances the persons were discussing an objective situation—the comparative number of geometrical figures on two cards or the veracity of a subject as manifested by his directly observable behavior. If anything the present experiment afforded a wider variety of items for discussion and would presumably be more sensitive to the sex variable. It seems doubtful if there is any appreciable sex difference in the ability to profit by discussion.

Accuracy of Judgments of Veracity

A consideration of either the initial or final judgments in Table I. indicates that, in this particular type of situation, estimates of veracity are very unreliable. Further analysis however reveals a rather general tendency. This is shown in Table II.

TABLE II

PER CENT. CORRECT JUDGMENTS (BOTH SEXES)

	Subject Telling Truth	Subject Lying
Before discussion	37	58
After discussion	33	62

The initial judgments when the subject is actually telling the truth are 37 per cent. correct and when he is lying 58 per cent. correct. The corresponding figures after discussion are 33 and 62 per cent. There is an obvious tendency to give a correct L judgment more frequently and a correct T judgment less frequently than the chance expectation of 50 per cent. This constant error is increased very slightly by the discussion. A more detailed table shows it to be slightly greater—5 or 6 per cent.—with men than with women. It appears that the situation creates an attitude of suspicion toward the subject and he is thus too frequently judged to be lying.

ABILITY TO JUDGE VERACITY OF ONE'S OWN SEX

Table III. gives the results bearing upon the ability to judge veracity of persons of one's own sex as compared with persons of the opposite sex.

TABLE III

PER CENT. CORRECT JUDGMENTS

	Women Judging Women	Women Judging Men	Men Judging Women	Men Judging Men
Before discussion	45	49	52	41
After discussion	52	51	52	38

The successive columns of the table indicate the per cent. of correct judgments made by female jurors with female subjects, female jurors with male subjects etc. The only marked divergence in the table from the chance 50 per cent. correct is in the greater inaccuracy of male jurors in judging men. A more detailed table shows that this is due primarily to the tendency, mentioned above, to give L judgments too frequently. Probably the male subjects in their efforts to appear guilty while telling the truth were more successful in deceiving others of their own sex. It is improbable that this indicates a fundamental sex difference.

Conclusions

In an experiment in which the subject lied or told the truth about an imaginary crime, a group of persons judged his veracity by his observable reactions during examination. After five minutes open discussion a second judgment was made. Approximately 250 judgments of this sort were available.

There was a considerable tendency to change one's decision as a result of the discussion but the change was in the wrong direction about as often as in the right. There was no appreciable sex difference in the ability to profit by discussion.

In the present situation there was a constant tendency to consider too frequently that the subject was lying. There was no clear difference in the ability to judge the veracity of persons of one's own as compared with those of the opposite sex.

PRACTICE EFFECTS IN INTELLIGENCE TESTS[1]

BY KNIGHT DUNLAP AND AGNES SNYDER

The Johns Hopkins University

The possible practice effects in intelligence tests are of importance where groups of reactors are to be given relative gradings on tests. If some of the reactors have had experience in the taking of tests; especially tests of the same general sort as that on which the grading is to be done; while others have had less experience, or none at all, with tests of the type used; it is obvious that the ratings will be unfair, unless no considerable practice effects can be demonstrated. The matter is of especial importance in connection with tests for admission to college or school, since the possibility of 'coaching' for these tests must be definitely known, or else the tests are misleading.

For the purpose of a preliminary investigation into practice effects, a college class of 44 men, mostly seniors, was tested four times, with intervals of approximately three weeks between the tests, using four forms of the Army "Alpha" composite test. The Alpha test was selected not because of any assumed superiority of that test over others, but because a number of nearly equivalent forms of that test were available, whereas there were at most only two forms of any other composite test available, with the exception of one standardized test for which the price asked was nearly a dollar each, rendering it prohibitive for experimental work.

The 'practice' in this investigation consisted solely in the taking of the tests: no attempt was made to give further practice between the tests, although this might have been done by using various single tests corresponding in type to the parts of the Alpha composite. No attempt has been made so far to find what effect practice on tests of one type

[1] A paper presented before the Southern Society for Philosophy and Psychology, New Orleans, April 23, 1920.

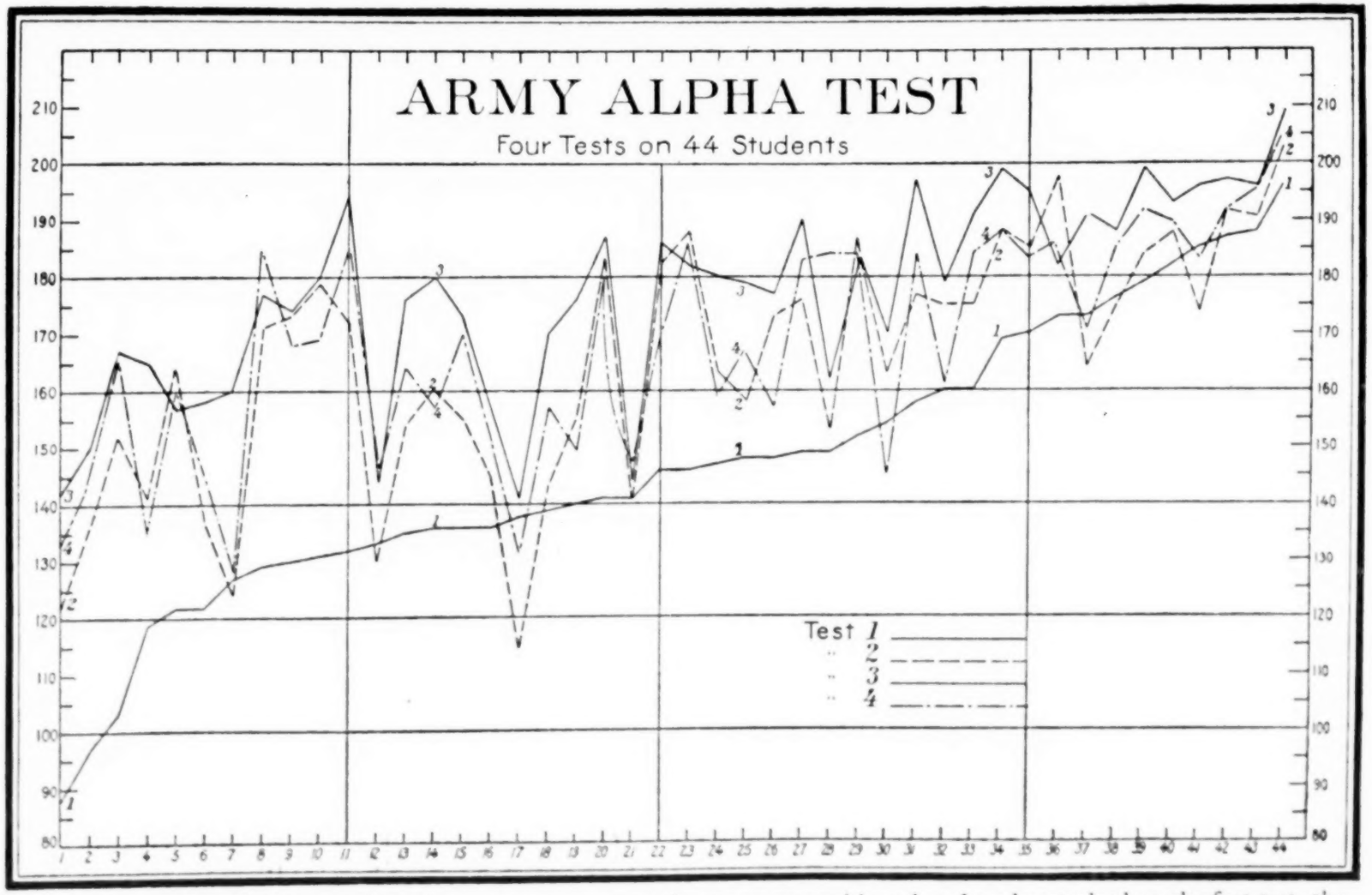


DIAGRAM 1. The Abscissæ numbers, from 1 to 44, represent the men, arranged in order of grades attained on the first test, the Graph for which is marked I. The Ordinate numbers, at the left, indicate the scale of markings in absolute scores on the test, not percentages. Graphs 2, 3 and 4 represent the grades received by the men on the second, third and fourth tests respectively.

exercises on tests of a different type. The purpose has been merely to obtain data under conditions which might most easily occur; namely, where persons expecting to submit to a certain sort of group-test might obtain previous forms of that test and study them or work through them for practice.

The results of the four tests are presented in Diagram 1. For this diagram, the members of the class were arranged in the order of their grades on the first test, from low (left) to high (right). The grades for the second, third and fourth tests were then plotted for the same arrangement. The reason for the specific arrangement is that thereby one fairly simple graph is obtained, making comparison with the graphs of the succeeding tests easy.

It will be noticed that there is a general improvement on the second test, and again on the third test, but a general falling back on the fourth. While five men on the second test fell below their marks on the first test, and one other made exactly the same mark on both tests; and five men on the third test fell below their marks on the second; every one made better marks on the third test than on the first. That is: not only was there a striking general improvement from the first to the second, and from the second to the third, but *every man* showed improvement on the third test, over the first.

The improvement on the second and third tests, and the slip-back on the fourth, is shown in another way in Diagram 2, where the ordinates represent the number of men making improvements, or losses, of magnitudes represented by the abscissae.

The main reason for the deterioration on the fourth test was made clear by the reports of the reactors. There was a practically unanimous report that the taking of the first, second and third tests was interesting; but that the fourth was a bore. The fact of lessened interest I inferred also from details of the behavior of the class during the test, before I obtained any reports. Of course, the fourth test should have shown at least as high an average grading as did the third if anything had depended on the test beyond the mere

personal ambition to make a good showing, which had been largely satisfied by the preceding tests.

It is probable that the four forms used were not exactly equivalent for this particular group of men. The most likely part in which differences would be found might be assumed to be the part on General Information; and the details of the average gains and losses on the different parts of the test (Diagram 3) indicate that the General Information

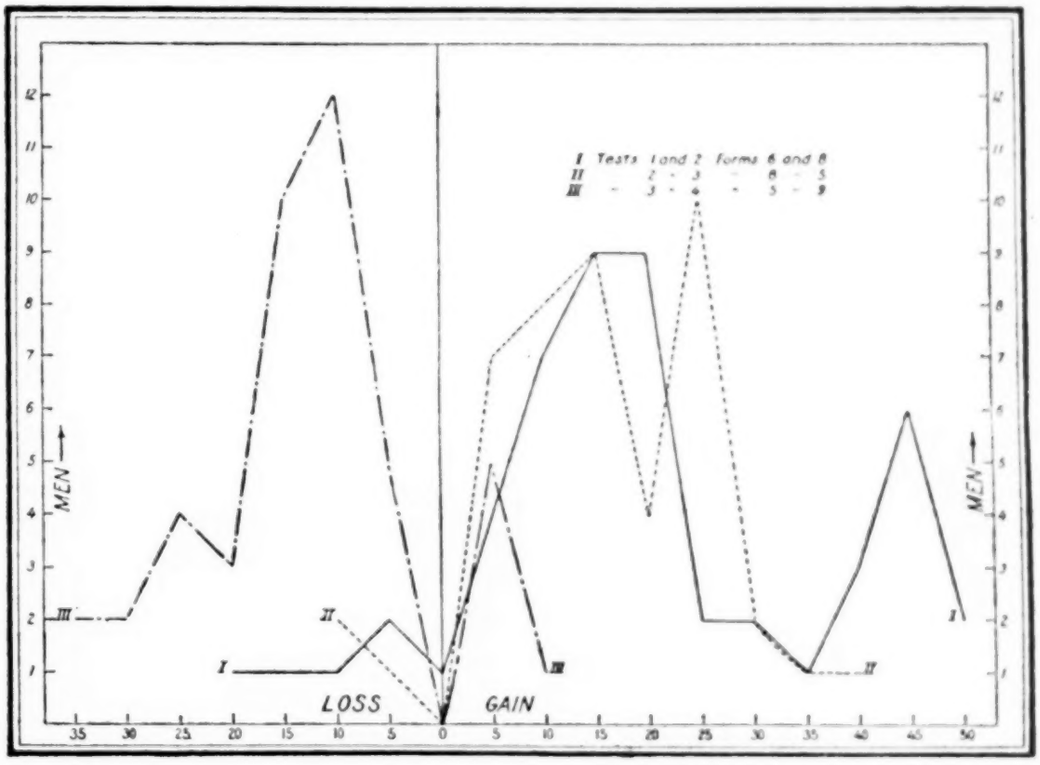


DIAGRAM 2. The abscissæ here represent the number of points of gain or loss between successive tests, the ordinates representing the number of men making a gain or loss not greater that the abscissa at the point of erection of the ordinate, but greater than the next smaller abscissal value marked. Graph I., for example, indicates that two men lost not more than five points on the second test (as compared with the first); one man lost more than five but not over ten; one lost not over fifteen, and one not over twenty.

part of the fourth test was indeed more difficult than the corresponding parts of the other forms. The probability of consequential differences in the other parts of the test is less definitely determinable; but since the excessive loss on the general information part of the fourth form accounts for only

a part of the general loss, the apathy or lessened interest of the class may be tentatively assumed as the most plausible major cause for the general decline on the fourth test.

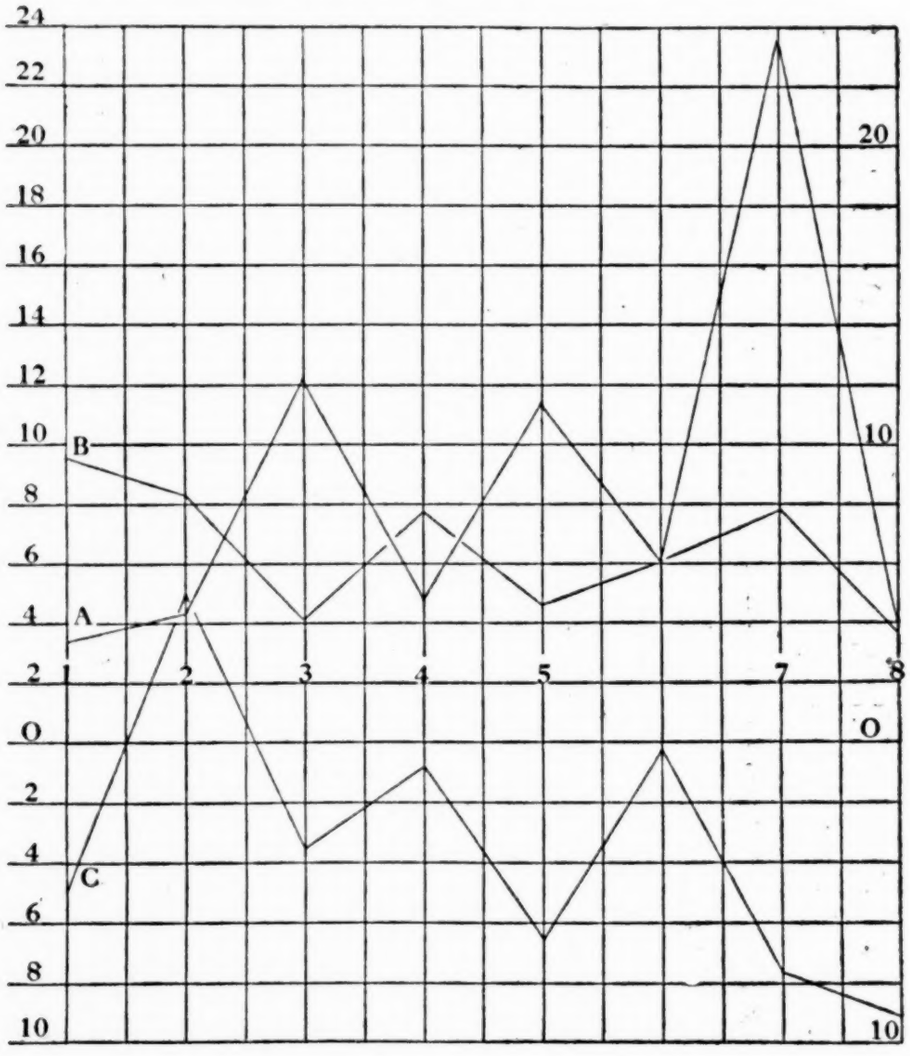


DIAGRAM 3. This represents the gains and losses, in percentages, on the eight parts of the composite test, the percentages being represented by the ordinates, and scaled at the left. The eight points marked on the horizontal axis represent the eight parts of the test. Graph *A* shows the average percentage of gains of the class on the second test as compared with the first: graph *B* represents the average percentage of gains on the third test as compared with the second; and graph *C* represents the gain and losses on the fourth test as compared with the third.

A detailed study of the gains and losses on the several parts of the test has given perplexing results. The close similarity of the corresponding parts of the four forms; especially the second part (problems in arithmetic), in which

the same wordings are used in some of the corresponding items of the several forms; and in the fourth part (number series), in which some identical number-series occur; may account for the significantly higher gain in these parts as compared with other parts of the tests. The problems of transfer of training are sharply presented here. The important question is whether repetition-forms of tests can be constructed in such a way as to test the same 'capacities,' and offer the same degrees of difficulty, without permitting a large amount of transfer. It is obvious that a considerable amount of experimental work must be done before tests can be used with assurance as to the absence of important practice effects.

Subject to the foregoing considerations, the following interesting points may be indicated:

In the first place, the relative rating on the third test is somewhat different from that on the first. If the tests had been used to select the best fourth of the group there would have been but little difference in the individuals selected, since 9 of the 11 who were in the first (high) quartile in the first test remain in the first quartile in the third, while the two who go back move back but to the next quartile. The tendency of movement backward or forward to adjacent quartiles is marked in the three lower quartiles. While this has no direct bearing on the present problem, it is a matter of passing interest to note the comparative uniformity in the results of the first and third tests in the first (high) quartile and the contrasting fluctuation of results in the fourth (low) quartile. Were the test given with a view to eliminating the lowest quartile the first and third tests would thus yield quite different results.

In the second place: if the men in the lower half of the class (as indicated by the first test) had had practice in the test equivalent to that actually obtained by the taking of the test twice, and the men who are in the upper half had not had practice, a single test on the class would have given a relative grading which would have differed markedly from either the first or the third of the actual ratings, and which (in all probability) would have been maximally unfair.

The practical conclusions to be drawn from such data are extremely important. The question as to which would be the fairest method of grading a group (of applicants for admission to college, let us suppose): grading them on a test with no preceding practice, or on a test preceded by uniform practice (the same amount for each individual), is of great theoretical interest, but of none at all practically, since groups will not be obtainable in either condition, if the system of testing becomes established and individuals know for months in advance that they are to be tested.

Obviously, if tests susceptible to practice effects are used, the only system which will be at all fair involves the condition that all candidates shall have old forms of the tests in advance, and shall have an opportunity to practice on them; or that in other ways all candidates shall receive an amount of practice which will put them in an equivalent "practiced" condition. The exact amount of practice required for that purpose must be experimentally determined, in order that the necessary minimum may be assured to all the candidates. Since coaching on such intelligence tests is distinctly possible and since the maximal unfairness is obtained where some are coached and others are not, pains must be taken to have all coached effectively.

Whether it is possible to develop types of test which are not susceptible to practice effects, and which yet test the same 'capacities,' and offer equal difficulties, is a question which cannot be answered definitely without prolonged research on that specific problem. The best conjectures we may make on this point are not encouraging. However, if the plan indicated above is followed systematically, we may be able to equalize practice effects and thus render them harmless.

TABLE I

CHANGES IN QUARTILES FROM TEST 1 TO TEST 3

	Fourth Quartile	Third Quartile	Second Quartile	First Quartile
To fourth	6	4	1	0
To third	3	4	4	0
To second	1	3	5	2
To first	1	0	1	9

TABLE II

AVERAGE PERCENTAGE OF IMPROVEMENT ON THIRD TEST OF MEN IN QUARTILES OF FIRST TEST

Fourth Quartile	Third Quartile	Second Quartile	First Quartile
54	11	40	18
53	41	33	12
64	44	29	11
46	37	31	25
36	21	41	11
35	3	13	8
33	31	31	10
46	36	16	19
44	3	39	19
62	36	19	13
General Average: 47.4	28.1	29.4	16

TABLE III

AVERAGE PERCENTAGE OF IMPROVEMENT ON THIRD TEST OF MEN IN QUARTILES OF THIRD TEST

Fourth Quartile	Third Quartile	Second Quartile	First Quartile
3	64	49	11
54	31	44	62
3	9	33	25
11	37	36	11
53	44	31	8
35	41	40	10
41	36	46	39
36	46	13	25
33	29	41	19
13	31	31	30
46	19	18	13
General Average: 28	35.2	34.7	23